

**Über die Beziehungen zwischen
Anspruchsniveau, Intelligenz und
Schulleistung (unter Berücksichtigung
der Stabilität und der Generalität des
Anspruchsniveaus)**

Inaugural-Dissertation

**im Fachbereich Philosophie,
Geschichte und Sozialwissenschaften
der Friedrich-Alexander-Universität
Erlangen-Nürnberg**

**Vorgelegt von
Christel-Astrid Brucks
aus Alfeld**



Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41548222_0057

Über die Beziehungen zwischen
Anspruchsniveau, Intelligenz und
Schulleistung (unter Berücksichtigung
der Stabilität und der Generalität des
Anspruchsniveaus)

Inaugural-Dissertation

im Fachbereich Philosophie,
Geschichte und Sozialwissenschaften
der Friedrich-Alexander-Universität
Erlangen-Nürnberg

Vorgelegt von
Christel-Astrid Brucks
aus Alfeld



Tag der mündlichen Prüfung: 26. 6. 1978

Dekan:	Univ.-Professor Dr. K.-H. Ruffmann
Referent:	Univ.-Professor Dr. W. Traxel
Korreferent:	Univ.-Professor Dr. W. Toman

Vorwort

Ich danke allen, die zum Entstehen dieser Arbeit beigetragen haben. Mein besonderer Dank gilt meinem Lehrer, Herrn Prof. Dr. W. Traxel, der die Untersuchung angeregt und in allen Stadien unterstützend begleitet hat. Ihm verdanke ich wertvolle Ratschläge, die er mir in vielen Gesprächen gegeben hat.

Wichtige Voraussetzungen dafür, daß die Untersuchung in der geplanten Form durchgeführt werden konnte, haben der Leiter der Gesamtschule Hollfeld, Herr Oberstudiendirektor O. Hofmann, und sein Stellvertreter, Herr A. Saam, geschaffen. Ohne ihre Hilfe wäre die Eingliederung der Versuche in den organisatorisch recht komplizierten Schulbetrieb einer Gesamtschule kaum möglich gewesen.

Weiter danke ich meinen Kollegen: Herrn Dipl.-Math. H. Algayer für fachkundige Hilfe bei der Erstellung von Rechenprogrammen zur Datenauswertung und Herrn Dr. D. Heller für viele kritische Hinweise und Anregungen.

Herzlicher Dank gilt Frau I. Goldfuß für die große Mühe, die sie sich beim Schreiben des Manuskripts gemacht hat.

Bayreuth, den 28. 2. 1978

Christel-Astrid Brucks

Inhaltsverzeichnis

	Seite
1. Einleitung	1
2. Zum Stand des Problems	3
2.1. Grundbegriffe der Anspruchs- niveauforschung	4
2.2. Untersuchungen zur Generalität und Stabilität des Anspruchs- niveaus	6
2.3. Neuere Forschungsansätze zur Leistungsmotivation	12
2.4. Anspruchsniveau, Intelligenz und Schulleistung	17
2.4.1. Zur Problematik des Intelligenz- begriffs im Hinblick auf die Schulleistung	17
2.4.2. Untersuchungen zur Beziehung zwischen Anspruchsniveau, Intelligenz und Schulleistung	18
2.5. Entwicklung der Fragestellung	32
3. Methodik	40
3.1. Entwicklung der Anspruchs- niveauverfahren	40
3.2. Hauptversuche	53
3.2.1. Auswahl der Stichprobe	54
3.2.2. Erfassung der Intelligenz und der Schulleistung	55
3.2.3. Durchführung der Versuche	61
3.2.3.1. Hauptversuche I	62
3.2.3.2. Hauptversuche II	75

4.	Darstellung und Diskussion der Ergebnisse	79
4.1.	Beschreibung der Ausgangsdaten	79
4.1.1.	Zur Variablen Intelligenz	88
4.1.2.	Zur Variablen Anspruchsniveau	90
4.1.3.	Zur Variablen Schulleistung	96
4.1.4.	Diskussion	102
4.2.	Zur Generalität des Anspruchsniveaus	108
4.2.1.	Diskussion	108
4.3.	Zur Stabilität des Anspruchsniveaus	113
4.3.1.	Diskussion	118
4.4.	Zur Beziehung zwischen den Variablen Anspruchsniveau, Intelligenz und Schulleistung	121
4.4.1.	Diskussion	128
5.	Zusammenfassung	140
	Literatur	144
	Anhang	150

1. Einleitung

Die Erforschung konstituierender Bedingungen für das Zustandekommen geistiger, speziell schulischer Leistungen hat sich bisher mehr auf die intellektuellen Voraussetzungen konzentriert. Der dynamische Aspekt der Leistung ist demgegenüber vernachlässigt worden. Sieht man von den inzwischen fast in Vergessenheit geratenen willenspsychologischen Ansätzen ab, so ist die Frage nach den motivationalen Anteilen der Schulleistung relativ lange als Untersuchungsgegenstand ausgespart worden.

Für diese Sachlage sind, wie ein historischer Rückblick deutlich macht, neben inhaltlichen Gründen (zum Beispiel der Einfluß des Behaviorismus, dessen restriktives Gegenstandskonzept die Untersuchung der Motivation als Erlebnisvariable streng genommen nicht zuließ) wohl insbesondere methodische Schwierigkeiten verantwortlich zu machen. Dieses Problem einer angemessenen methodischen Umsetzung erscheint nach wie vor unbefriedigend gelöst.

Im Konzept des Anspruchsniveaus bietet sich ein methodischer Weg an, dessen entscheidender Vorzug darin gesehen werden kann, einen phänomennahen Zugang zur Leistungsmotivation (verstanden als verpflichtend erlebter Anspruch an die eigene Leistung) zu ermöglichen.

Ein erstes Ziel der vorliegenden Untersuchung besteht darin, die Beziehung zwischen Anspruchsniveau¹⁾, Intelligenz und Schulleistung näher zu erkunden. Die Problematik der Messung der Intelligenz und der Schullei-

1) Anspruchsniveau wird im weiteren mit "AN" abgekürzt.

stung wird dabei (als bereits hinlänglich behandelt) zugunsten des ANs zurückgestellt.

Um die in Frage stehende Beziehung sinnvoll untersuchen zu können, wird vorab die Frage nach der intra-individuellen Konsistenz des ANs über die Zeit (Stabilität) und über verschiedene Leistungen (Generalität) zu klären versucht. - Wenn das AN sich als eine aussagekräftige Variable erweisen soll, dann muß sie - durchaus in der Art des Intelligenzmaßes - übergreifend, weniger situations- bzw. leistungsspezifisch sein, zugleich aber auch eine gewisse Stabilität, das heißt Merkmalskonstanz aufweisen.

Um hinsichtlich der Intelligenz und der Schulleistung einen möglichst umfassenden Variationsbereich einbeziehen zu können, wird die Untersuchung an Schülern einer Gesamtschule ausgeführt.

Für die Betrachtung der Generalität des ANs werden in bezug auf die jeweils abverlangte Leistung unterschiedliche AN-Verfahren eingesetzt. Zur Kontrolle der Stabilität des ANs wird für die AN-Verfahren eine Versuchswiederholung vorgenommen. Das Zeitintervall zwischen beiden Messungen wird relativ lang (3 1/2 Monate) bemessen, um etwaige Merkmalsschwankungen tatsächlich zur Wirkung gelangen zu lassen. Eine hinreichende Klärung steht für beide Aspekte der Konsistenz des ANs noch aus. Eine Kontrolle der Stabilität, die eine längere Zeitspanne einschließt, fehlte bislang.

2. Zum Stand des Problems

Das Problem der motivationalen Anteile der Leistung und das Problem eines angemessenen methodischen Zugangs bilden eine Frage, die zunächst lange Zeit nicht gestellt wurde und deren Schwierigkeiten in der methodischen Umsetzung man bisher nicht befriedigend zu lösen vermochte. Damit ist der Hintergrund für den derzeitigen Stand der Forschung in kurzer Form skizziert.

Historisch betrachtet waren es neben inhaltlichen Gründen wohl nicht zuletzt methodische Schwierigkeiten, die die Entwicklung einer spezifisch psychologischen Motivationsforschung erschwert haben. Diese Gründe sollen im folgenden nur kurz angesprochen und später (siehe unter Punkt 2.5.) noch ausführlicher dargestellt werden.

Die Vorstellungen über "Motivation" waren zunächst über mehrere Jahrzehnte davon beeinflusst, daß in diesem Bereich entweder das tierische Verhalten oder das pathologische Verhalten des Menschen als bevorzugter Forschungsgegenstand galt. Darüber hinaus ist es wohl vornehmlich dem restriktiven Forschungsansatz des Behaviorismus zuzuschreiben, daß die Motivation überwiegend auf ihren physiologischen Aspekt reduziert oder teilweise in ihrer Bedeutung überhaupt negiert wurde. Zwar haben sich die Vertreter der tiefenpsychologisch orientierten Psychologie bereits vor diesen behavioristischen Restriktionen intensiver mit Fragen der menschlichen Antriebe befaßt, doch standen hier klinisch-therapeutische Belange im Vordergrund. Schließlich sind wesentliche Erschwernisse für die Untersuchung des Motivationsgeschehens in Problemen der methodischen Handhabung zu sehen.

In der allgemeinen Psychologie setzte LEWIN einen Anfang (sieht man von der älteren Willenspsychologie ab). Die Motivation nimmt in der experimentellen und theoretischen Arbeit LEWINs eine zentrale Stellung ein. Mit seiner dynamischen und spezifisch psychologischen, vom jeweiligen Erleben ausgehenden Betrachtungsweise wurden Möglichkeiten für die Analyse des Zustandekommens von Leistungen eröffnet. Insbesondere in den Untersuchungen über das Anspruchsniveau wurden hierzu wesentliche Grundlagen erarbeitet.

2.1. Grundbegriffe der Anspruchsniveauforschung

Im folgenden soll ein komprimierter Überblick über die AN-Forschung gegeben werden.¹⁾ Es werden grundlegende Begriffe eingeführt und die wichtigsten Ergebnisse dargestellt.

Motivationspsychologisch können wir das Handeln unter drei Aspekten fassen: Wahl (zwischen möglichen Zielalternativen), Intensität und Ausdauer. In der experimentellen Motivationsforschung sind dies die gebräuchlichsten Verhaltenskriterien.

In der AN-Forschung hat sich das Kriterium der Wahl in besonderem Maß bewährt. Im AN-Versuch erhält die Vp Gelegenheit, sich vor der (mehrfach) wiederholten Lösung von Aufgaben - mit abstufbarem Schwierigkeitsgrad - Ziele zu setzen. Dieses experimentelle Grundmuster wurde

1) Die von LEWIN ausgehende, außerordentlich anregende Wirkung auf die Erforschung des ANs läßt sich daran ermessen, daß bis heute etwa 500 (vorwiegend amerikanische) Arbeiten zum AN vorliegen.

1930 in der klassischen Arbeit zum AN von F. HOPPE, einem Schüler LEWINs, eingeführt.

Unter AN ist nach HOPPE (1930) "die Gesamtheit ... (der) ... mit jeder Leistung sich verschiebenden, bald unbestimmteren, bald präziseren Erwartungen, Zielsetzungen oder Ansprüche an die zukünftige eigene Leistung" (S. 10) zu verstehen. Diese Definition umfaßt nicht nur explizite, sondern auch implizite Zielsetzungen. Mit der von FRANK (1935) eingeführten Quantifizierung des ANs nahm man eine Restriktion des Begriffes im Interesse einer methodischen Präzisierung in Kauf.

Die wichtigsten Ergebnisse der AN-Forschung lassen sich auf das folgende zusammenfassen:¹⁾

Die Dynamik des ANs ist anhand von drei Arten von Bezugssystemwirkungen wiederholt nachgewiesen worden:

1. Das Auftreten von Erfolgs- und Mißerfolgserlebnissen ist nicht fest an eine bestimmte Leistungshöhe bzw. -menge gebunden. Entscheidend für das Erleben von Erfolg und Mißerfolg ist das Verhältnis der erreichten zur angestrebten Leistung. 2. Nur innerhalb eines mittleren subjektiven Schwierigkeitsbereiches werden Erfolg und Mißerfolg erlebt. Das AN bildet sich nur in bezug auf Ziele einer bestimmten, eng umschriebenen Schwierigkeitszone heraus. 3. Ein ebenfalls häufig repliziertes Ergebnis sind die "Verschiebungsgesetzmäßigkeiten" des ANs. In der Regel wird das AN nach Erfolg herauf- und nach Mißerfolg herabgesetzt.

1) Vgl. im einzelnen LEWIN et al. 1944; HECKHAUSEN 1965.

Neben diesen Bezugssystemwirkungen, zu denen innerhalb der AN-Forschung die meisten Untersuchungen unternommen wurden, gibt es weitere, persönliche und situative Faktoren, die die inter- und intraindividuelle Variation des ANs beeinflussen.

Die Fragen der Generalität und der Stabilität sollen etwas eingehender behandelt werden.

2.2. Untersuchungen zur Generalität und Stabilität des Anspruchs-niveaus

Ausgehend von einer - immer wieder festgestellten - beachtlichen interindividuellen Variabilität des ANs, erhebt sich die Frage, welche Resultate hier die intra-individuelle Betrachtung erbringt. - Die Auseinandersetzung mit diesem Problem führt zu den Grundlagen, deren Klärung vorauszusetzen ist, um überhaupt mit dem Konzept des ANs sinnvoll arbeiten zu können.

Intraindividuelle Unterschiede bzw. Übereinstimmungen des ANs sind in zweierlei Hinsicht zu bestimmen: die Konsistenz über die Zeit (Stabilität) und über verschiedene Leistungen (Generalität).

Die Stabilität des ANs ist bisher nur vereinzelt untersucht worden.

FRANK, ein Schüler LEWINS, legte 1935 die erste Untersuchung zur Stabilität des ANs vor. Es wurden drei AN-Aufgaben verwendet (Drucken von Wörtern, Erfassen räumlicher Beziehungen, Ringwerfen), die sich in bezug auf die abverlangte Leistung (intellektuell, motorisch)

und das zugrunde gelegte Leistungsmaß (Schnelligkeit, Genauigkeit) unterschieden. Drei Versuchsgruppen zu je 12 Vpn (Studenten) wurden jeweils zwei der drei AN-Aufgaben im Einzelversuch vorgelegt. Nach einer Woche wurde eine Versuchswiederholung durchgeführt. Die mitgeteilten Retest-Koeffizienten¹⁾ sind mit Werten zwischen .57 und .75 statistisch gesichert; lediglich beim Ringewerfen ergibt sich ein nicht signifikanter Wert (.26), wofür FRANK neben Mängeln in der Versuchsdurchführung den mehr spielerischen Charakter des Ringewerfens verantwortlich macht.

Die von FRANK festgestellte Stabilität bezieht sich auf ein Zeitintervall von einer Woche. In diesem relativ kurzen Intervall mag sich ein Merkmal durchaus noch als stabil erweisen, das über eine längere Zeitspanne instabil sein kann, abgesehen davon, daß die Ergebnisse auf einer verhältnismäßig kleinen Datenmenge (N = 12) basieren.

ROTTER (1942) betrachtete die Stabilität des ANs mit einer Versuchswiederholung nach vier Wochen. Im Einzelversuch wurden 32 Vpn (Krankenhausangestellte mit einem Durchschnittsalter von 25 Jahren) mit einer relativ einfachen motorischen Aufgabe, dem "ROTTER Aspiration Board", untersucht. ROTTER erhält für die Zieldiskrepanzen²⁾ aus der 1. und 2. Messung einen Stabilitätskoeffizienten von .32.³⁾

1) Berechnung von Rangkorrelationen für die Zieldiskrepanzwerte (mittlere Differenz zwischen jeweils erreichter Leistung und nachfolgender AN-Setzung) aus der 1. und 2. Messung.

2) Zieldiskrepanz wird mit "ZD" abgekürzt.

3) Zur Art der berechneten Korrelation und zur Signifikanz werden keine Angaben gemacht.

Die Frage, ob und in welchem Ausmaß nach einem Zeitintervall von einem Monat noch Erinnerungseinflüsse anzunehmen sind, bleibt offen.

In einer Untersuchung an vier- bis sechsjährigen Kindern (BRUCKS 1977) wurden Stabilitätsschätzungen vorgenommen, die ein Zeitintervall von 19 Wochen einschließen. Das AN von 54 Vpn wurde im Einzelversuch anhand einer einfachen motorischen Aufgabe (Bälle in einen Trichter werfen) erfaßt. Hierzu wurde eine Versuchsanordnung eingesetzt, die eine - angesichts des Alters der Kinder notwendige - Veranschaulichung der in bezug auf das AN wichtigen Informationen (jeweils Leistungsvoraussetzung, Leistungsvornahme und Leistungsergebnis) ermöglicht. Es ergab sich ein Rangkorrelationskoeffizient von .54 (sign. bei 0,1 %) für die ZD-Werte der ersten und zweiten Messung.

Diese Ergebnisse, die sich auf eine vergleichsweise lange Zeitspanne stützen, weisen bereits auf eine gewisse Merkmalskonstanz des ANs hin.

Die Frage nach der intraindividuellen Konsistenz des ANs hinsichtlich verschiedener Leistungen wird bereits in der Originalarbeit zum AN (HOPPE 1930) gestellt und auch berücksichtigt. HOPPE wählte "absichtlich Handlungen von möglichst verschiedenem Typus, damit auch etwaige Abhängigkeiten dieses Vorgangs von der Handlungsstruktur in Erscheinung treten" (S. 12). Die qualitativen Befunde HOPPEs geben erste, vorläufige Hinweise zur Konsistenz, denen FRANK als erster gezielt nachgeht.

FRANK (1935) unternahm damit erstmalig den Versuch, die Generalität des ANs quantitativ zu bestimmen. Die Anlage des Versuchs sei - zur verkürzten Beschreibung¹⁾ - schematisch dargestellt:

Versuchs- gruppe	A N - V e r f a h r e n (Leistung/Leistungsmaß)		
	Drucken v. Wörtern (motorisch/ Schnellig- keit)	Erfassen räumlicher Beziehungen (intell./ Schnellig- keit)	Ringe- werfen (motorisch/ Genauig- keit)
I (N=12)	x ⁺)	x	
II (N=12)	x		x
III (N=12)		x	x

⁺) x = jeweils durchgeführtes AN-Verfahren

Die Vpn wurden einzeln untersucht; nach einer Woche fand eine Versuchswiederholung statt. Die über die Zieldiskrepanzen errechneten Rangkorrelationen ergeben Werte zwischen -.08 (Erfassen räumlicher Beziehungen/Ringewerfen) und +.70 (Drucken von Wörtern/Ringewerfen). - Eine eindeutige Interpretation der Ergebnisse im Sinn der für die AN-Verfahren eingeführten Unterschiede (abverlangte Leistung/Leistungsmaß) erscheint äußerst schwierig. Zudem ist zu bedenken, daß die deutlich differierenden Koeffizienten sich jeweils nur auf die Daten von 12 Vpn stützen.

1) Zur Beschreibung des Versuchs siehe auch S. 6 f.

FRANK sieht in seinen Ergebnissen eine Bestätigung dafür, "that there are certain traits of the personality, measured through the behavior of the level of aspiration, which are to a large extent independent of the physical nature of the situations in which they are aroused" (1935, S. 127).

Diese Annahme unterzog GOULD (1938, 1939) einer kritischen Überprüfung, indem sie eine größere Stichprobe (N = 82 männliche Studenten) wählte und eine größere Zahl verschiedener AN-Verfahren verwendete. Es wurden im Einzelversuch sechs verschiedene AN-Aufgaben vorgegeben: 1. "words" (Synonyma finden); 2. "steadiness" (motorische Aufgabe zur Prüfung der Ausdauer); 3. "addition" (Addition von je vier einstelligen Zahlen); 4. "symbol-digit" (Zuordnung von Zahlen zu Symbolen); 5. "cancellation" (Durchstreichen bestimmter Buchstabenkombinationen); 6. "target" (Zielwerfen auf eine Scheibe). Mit Ausnahme des "target" war für alle Aufgaben eine Zeitbegrenzung festgelegt. Die in bezug auf die Leistung unkorrelierten Aufgaben waren auf zwei Versuchssitzungen mit einem durchschnittlichen Zeitabstand von einer Woche verteilt. Die über die Zieldiskrepanzwerte errechneten Interkorrelationen zwischen den einzelnen AN-Verfahren streuen zwischen .04 ("steadiness"/"target") und .44 ("addition"/"symbol-digit") mit einer mittleren Korrelation von .29. Die Koeffizienten fallen damit gegenüber den von FRANK (1935) mitgeteilten Werten deutlich geringer aus. GOULD sieht in diesen niedrigen Interkorrelationen keine Bestätigung der - von FRANK angenommenen - Generalität des ANs, unter der Bedingung, daß das AN über das Zieldiskrepanz-Maß erfaßt und an Leistungen untersucht wird, die nicht bzw. nur geringfügig miteinander korrelieren. Es werden weitere Faktoren in ihrem möglichen Einfluß auf die Generalität

diskutiert. Insbesondere weist GOULD auf Schwierigkeiten hin, die aus der Verwendung des Zieldiskrepanz-Maßes entstehen: für einen Prozeß (das AN) wird ein Maß (die ZD) herangezogen, in dem sich lediglich das Endresultat eines solchen Prozesses abbilden kann. - Auf die Frage der Angemessenheit des ZD-Maßes werden wir an späterer Stelle (siehe S. 134 ff.) näher eingehen. -

Die späteren Untersuchungen zur Generalität des ANs und den Bedingungen ihres Auftretens (GARDNER 1939; HEATHERS 1942; PRESTON und BAYTON 1942; KOWALEWSKY 1955) zeigen in der methodischen Handhabung ein allen gemeinsames Prinzip. Man versucht, über verstärkte Kontrolle bzw. gezielte Variation der für die Generalität als wichtig angenommenen Bedingungen zu präzisieren Aussagen zu gelangen. Dabei bedient man sich bevorzugt eines Ansatzes, der in zweierlei Hinsicht problematisch erscheint: Mit dem Ziel, experimentell Erfolg bzw. Mißerfolg zu induzieren, werden der Vp im AN-Versuch nicht ihre tatsächlich erreichten Leistungen mitgeteilt, sondern fiktive Leistungswerte, die einer im vorhinein festgelegten Leistungskurve folgen. Dabei wird ungeprüft vorausgesetzt, daß die beabsichtigte der tatsächlichen Wirkung entspricht, was aber - nach allem, was man über die Bezugssystemwirkung des ANs weiß (vgl. S. 5) - keineswegs als zwingend angenommen werden kann. Überdies bleibt es zweifelhaft, in welchem Maß es gelungen ist, die vorgegebenen Leistungskurven, die zum Teil recht atypisch verlaufen (z. B. Kurventeile mit abruptem Abfall und Anstieg), glaubhaft zu gestalten.

Die Frage, inwieweit in diesen Untersuchungen mit Artefakten gerechnet werden muß, läßt sich nicht beantworten.

Zusammenfassend ist festzustellen, daß die bisherigen Arbeiten zur Generalität und zur Stabilität des ANs noch keine befriedigende Klärung erbracht haben.

In der vorliegenden Arbeit sollen daher die beiden Aspekte der intraindividuellen Konsistenz des ANs weiter verfolgt werden.

2.3. Neuere Forschungsansätze zur Leistungsmotivation

Seit den Arbeiten von LEWIN und seinen Schülern hat die Erforschung der Leistungsmotivation in zunehmendem Maß Beachtung gefunden. Der Ansatz LEWINS wurde - mit dessen Emigration - von Deutschland nach Amerika getragen.

Auf der Suche nach brauchbaren Meßverfahren der Leistungsmotivation verfolgte man neue Wege. Das Konzept des ANs geriet dabei offensichtlich weitgehend in Vergessenheit.¹⁾

Im wesentlichen sind es zwei Verfahren bzw. Verfahrensgruppen, über die man einen Zugang zur Leistungsmotivation suchte: die indirekten projektiven Verfahren des TAT (Thematischer Auffassungs-Test) und die Fragebogenmethoden.

Die fundamentale Annahme - und zugleich die Rechtfertigung - des projektiven Ansatzes besteht darin, daß das erlebende und handelnde Individuum nicht in der Lage sei, über seine Leistungsmotivation unmittelbare Aussagen zu machen. "Motivation ... ist offensichtlich zu

1) In Darstellungen der Methoden zur Messung der Leistungsmotivation (z. B. KRUMBOLTZ 1957; HECKHAUSEN 1977) wird das AN vielfach gar nicht erwähnt.

abstrakt und erlebensjenseitig, um darüber Rechenschaft geben zu können." (HECKHAUSEN 1968, S. 639). Folglich - so die Befürworter des TAT - müsse man einen mittelbaren Weg über projektive Phantasie wählen, um die Leistungsmotivation möglichst unverfälscht zu erfassen. Das heißt, der indirekte Zugriff wird zugleich als der direkteste angesehen. Der TAT, ein inhaltsanalytisches Verfahren, gilt hier als die Methode der Wahl.

McCLELLAND und seine Mitarbeiter haben den auf MURRAY (1938, 1942) zurückgehenden TAT adaptiert und in einer standardisierten Form zur Messung der Leistungsmotivation eingesetzt (McCLELLAND et al. 1953).

HECKHAUSEN (1963 a) stellt eine Verfeinerung der McCLELLANDSchen TAT-Version vor, die nicht nur Aussagen über die Intensität der Leistungsmotivation, sondern auch über deren Richtung, die beiden Tendenzen "Hoffnung auf Erfolg" und "Furcht vor Mißerfolg", ermöglicht.

Schließlich haben MEYER, HECKHAUSEN und KEMMLER (1965) noch eine Kinder-Version dieses Verfahrens entwickelt.

Auf die grundsätzliche Problematik der Projektionsannahme soll hier nur hingewiesen werden (ausführlich behandelt z. B. bei HÖRMANN 1964; KORNADT 1964).

Die Kritik am TAT (einschließlich der verschiedenen Verfahrensvarianten) richtet sich insbesondere auf seine psychometrischen Güteeigenschaften.

Der TAT besitzt nur eine sehr geringe innere Konsistenz (vgl. HECKHAUSEN 1963 a, S. 77). Nachdem für den TAT in Anspruch genommen wird, Leistungsmotivation als relativ überdauernde Größe zu messen, kommt der Wiederholungszu-

verlässigkeit eine besondere Bedeutung zu. Die zur Retest reliabilität vorliegenden Schätzungen sind hingegen kaum geeignet, diese Annahme zu stützen. Auch bei sorgfältiger Kontrolle der äußeren Versuchsbedingungen (Durchführung zur selben Tageszeit, am selben Ort, mit demselben V1) liegen die erzielten Wiederholungskorrelationen nicht höher als zwischen .41 und .51¹⁾ (siehe HECKHAUSEN 1963 a, S. 79). Sättigungseffekte sowie die Instruktion, beim Geschichtenerzählen möglichst viel Phantasie zu entfalten (vgl. HECKHAUSEN 1963 a, S. 58 f.), wirken sich reliabilitätsmindernd aus. Die Anfälligkeit des TAT gegenüber Variationen in den Versuchsbedingungen bedeutet insofern einen schwerwiegenden Nachteil, als sie auch bei einer ersten Versuchsdurchführung wirksam werden und damit die Zuverlässigkeit reduzieren kann. Eine weitere Quelle der Fehlerstreuung liegt in der unvollkommenen Objektivität der Auswertung. Korrelationen zwischen zwei "unabhängigen"²⁾ Beurteilern erreichen Werte zwischen .81 und .89 (HECKHAUSEN 1963 a, S. 83).

Vor dem Hintergrund der theoretischen Zusammenhänge zwischen Objektivität, Reliabilität und Validität ergeben sich eher ungünstige Bedingungen für die in Frage stehenden Gültigkeitsnachweise des TAT.

Die Validierungsbefunde zum TAT stellen sich, insgesamt betrachtet, als sehr uneinheitlich, vielfach widersprüchlich dar, was häufig zu kontroversen Beurteilungen der

-
- 1) Bei Ausscheiden einer Vp ergeben sich Koeffizienten zwischen .42 und .59.
 - 2) Der Autor weist darauf hin, daß Übereinstimmungen in dieser Höhe "wohl nur erreicht werden, wenn beide Beurteiler sich gründlich in den Inhaltsschlüssel eingearbeitet und während dieser Zeit ihre Verrechnungen gelegentlich miteinander verglichen haben" (HECKHAUSEN 1963 a, S. 82).

Sachlage führt (vgl. etwa HECKHAUSEN 1963 a; KLINGER 1966; SCHMALT 1976 a; WASNA 1976). Der Gültigkeitsbereich, den der TAT nachweislich abzudecken vermag, ist also nach wie vor unzureichend geklärt, so daß eindeutige Aussagen gegenwärtig noch nicht möglich sind.

Neben dem auf einem indirekten Meßkonzept basierenden TAT-Verfahren, dessen psychometrische Eigenschaften sich als wenig befriedigend erwiesen haben, gibt es Versuche, die Leistungsmotivation über direkt arbeitende Methoden, das heißt Fragebogen zu erfassen. Mit diesen Fragebogeninstrumenten sollte ein Ersatz für das inhaltsanalytische Verfahren geschaffen werden, weil dieses eine zeitaufwendige Auswertung (mit entsprechender Einarbeitung) verlangt und überdies hinsichtlich seiner internen Konsistenz und seiner Wiederholungszuverlässigkeit nicht den zu stellenden Anforderungen genügt.

Die Fragebogenverfahren zeigten zwar eine höhere Reliabilität, doch ihre zunächst erwartete Überlegenheit gegenüber dem TAT, namentlich im Hinblick auf die Validität und die Ökonomie, hat sich nicht bestätigt. - Eine Übersicht über die entwickelten Fragebogen einschließlich der zur Reliabilität und zur Validität jeweils vorliegenden Befunde findet sich bei SCHMALT (1976 b).

Darüber hinaus wird häufig die Angemessenheit direkter Verfahren grundsätzlich in Frage gestellt. Diese, etwa von McCLELLAND (1958, 1971) und HECKHAUSEN (1967, 1968) vertretene Position wird in der Hauptsache damit begründet, daß die Vp nicht in der Lage sei, unmittelbare Aussagen über ihre Leistungsmotivation zu machen (vgl. dazu S. 12 f.), und daß sie zudem geneigt sei, ihre Antworten im Sinne des sozial Erwünschten zu verfälschen.

Zusammenfassend bleibt festzustellen, daß sowohl die TAT-Verfahren als auch die Fragebogenverfahren mit - jeweils spezifischen - Mängeln belastet sind, die zu einer nur wenig befriedigenden Messung der Leistungsmotivation führen.

Innerhalb der mit Beginn der fünfziger Jahre einsetzenden Leistungsmotivationsforschung McCLELLANDscher Provenienz, die im deutschen Sprachraum entscheidend durch HECKHAUSEN weitergeführt wurde, nimmt das TAT-Verfahren eine gewisse "Monopolstellung" ein. Der TAT wurde zum methodischen Ausgangspunkt, vor allem aber zur methodischen Grundlage einer ebenso intensiven wie extensiven Forschung. Die Entwicklung des TAT wird von seinen Vertretern als Durchbruch oder sogar als der eigentliche Beginn der Leistungsmotivationsforschung verstanden: "Seit ihren Anfängen hat die Leistungsmotivationsforschung einem indirekten - 'projektiven' - Zugriff vertraut" (SCHMALT 1976 a, S. 165).

Die anregende Wirkung des TAT auf die Forschung war un-
gemein stark. Der heuristische Wert dieses methodischen Ansatzes steht außer Frage. Problematisch ist jedoch die nahezu ausschließliche Verwendung eines Meßinstruments - zumal die psychometrischen Qualitäten dieses Verfahrens umstritten sind und bekanntlich die Ergebnisse nicht "besser" sein können als die Methode, mit der sie gewonnen wurden.

Diese Überlegungen führen zu der Frage, ob und in welchem Ausmaß bei den mit dem TAT gewonnenen Ergebnissen mit methodischen Abhängigkeiten bzw. Artefakten gerechnet werden muß.

In der vorliegenden Arbeit wird der Versuch unternommen, die Leistungsmotivation nicht über das Vehikel der Projektion, sondern über konkrete Anspruchsniveausetzungen zu fassen. (Näheres siehe unter Punkt 2.5., S. 32 ff.)

2.4. Anspruchsniveau, Intelligenz und Schulleistung

2.4.1. Zur Problematik des Intelligenzbegriffs im Hinblick auf die Schulleistung

Das Zustandekommen geistiger Leistungen wurde lange Zeit vorrangig unter dem Aspekt kognitiver Bedingungen betrachtet und auch untersucht. So wurden auch schulische Leistungsunterschiede weitgehend auf Intelligenzunterschiede zurückgeführt. Die Intelligenzunterschiede wiederum wurden vorwiegend erbbiologisch erklärt. Die Schulleistung galt also primär als eine Funktion der Intelligenz, deren interindividuell variierende Ausprägung genetisch bestimmt sei.

In der Tradition dieses Ansatzes, von dessen Angemessenheit und Ergiebigkeit man lange überzeugt war, wurden zahlreiche Untersuchungen über den Zusammenhang zwischen Intelligenz und Schulleistung angestellt. Die Fülle dieser Arbeiten ist inzwischen kaum noch zu überblicken. - Übersichten und weitere Hinweise geben z. B. FIPPINGER (1966), INGENKAMP (1975, S. 96 ff.), LAVIN (1965) und LÖSCHENKOHL (1973).

Zieht man ein Résumé aus den einschlägigen Untersuchungen, so ergeben sich - je nach der Stichprobengröße und -zusammensetzung und der Art der Erhebung und korrela-

tionsstatistischen Verrechnung der Intelligenz- und Schulleistungsdaten - unterschiedliche, im Durchschnitt mittelhohe Korrelationen (das heißt Koeffizienten um .5) zwischen den beiden Variablen. Dies bedeutet, daß durch Unterschiede der Intelligenztestleistungen nur ca. 25 % der Schulleistungsvariation erklärt werden können. Die Intelligenz erweist sich somit zwar als der relativ erklärungsmächtigste Faktor schulischer Leistung, zugleich wird aber auch die begrenzte Tragfähigkeit dieses Ansatzes deutlich.

Ogleich man die Rolle des "Fleißes" wohl nicht in Zweifel zog, wurden doch Unterschiede der Schulleistung vorwiegend auf Unterschiede der Intelligenz und weiter auf Unterschiede des genetischen Potentials zurückgeführt. Die Probleme, die eine solche, von theoretischen Voreingenommenheiten belastete "Schlußkette" aufwirft, bestehen offensichtlich nicht zuletzt darin, anstehende Fragen eher auszuschließen bzw. zuzudecken statt sie einer Lösung näherzubringen.

2.4.2. Untersuchungen zur Beziehung zwischen Anspruchsniveau, Intelligenz und Schulleistung

Solange man Intelligenz und Begabung als weitgehend erblich determinierte Potenzen begriff und diesen die vorrangige Erklärungslast schulischer Leistungsunterschiede zuwies - das heißt beobachtbare Leistungsdifferenzen nahezu für naturnotwendig und unveränderbar hielt -, wurde die Frage nach motivationalen Faktoren nicht zu einem untersuchenswerten Problem.

Minderleistungen trotz hoher Intelligenz wurden als Abweichung vom Normalen verstanden und mit affektiver

Leistungshemmung, neurotischer Leistungsblockierung oder der sogenannten Pseudodebilität erklärt (vgl. z. B. DÜHRSEN 1954; MÜLLER 1955; PIECHOCKI 1952; STAABS 1955; STERN 1956; ZULLIGER 1957). - Diese kasuistischen Studien mündeten indessen nicht in eine planmäßige Erforschung des Problems. Dies mag mit dem klinischen Ansatzpunkt und mit der Beschränkung auf die negative Variante der Leistungsabweichung zusammenhängen.

Verfolgt man die weitere Entwicklung, so scheinen mittlerweile zumindest die Schwierigkeiten, die zu Lasten bestimmter Voreingenommenheiten (siehe dazu auch das unter Punkt 2.4.1. Gesagte) gehen, weitgehend überwunden zu sein.

Damit sind aber zunächst nur die Voraussetzungen gegeben, die Frage nach den motivationalen Anteilen der Leistung überhaupt zu stellen. Nach wie vor problematisch ist die methodische Seite, das heißt eine adäquate Erfassung der Leistungsmotivation. - Die mit dem TAT und mit Fragebogenverfahren unternommenen Ansätze haben zu widersprüchlichen, in ihrer Aussagekraft zweifelhaften Ergebnissen geführt (z. B. MEYER et al. 1965; WASNA 1976) und sich als wenig erfolgreich erwiesen. (Zur Kritik der beiden Verfahren siehe unter Punkt 2.3.)¹⁾

Es bleibt zu fragen, inwieweit hier das Konzept des ANS weiterführend sein kann.

1) Auf die Darstellung von Untersuchungen auf der Grundlage dieser beiden Ansätze soll daher im folgenden verzichtet werden.

Im folgenden werden die einschlägigen Untersuchungen zur Beziehung zwischen AN, Intelligenz und Schulleistung dargestellt.

Pauline S. SEARS (1940, 1941) hat als erste Extremgruppen guter und schlechter Schüler in ihrer AN-Setzung verglichen. Sie untersuchte 36 Schüler des 4., 5. und 6. Schuljahres im Alter von 9 bis 12 Jahren. Nach Maßgabe vorausgegangener Schulleistungen (Noten, Testergebnisse) wurden drei in bezug auf Alter, Intelligenz und Verteilung der Geschlechter parallele Versuchsgruppen zu je 12 Vpn gebildet: eine "success"-Gruppe (gute Leistungen in allen Fächern), eine "failure"-Gruppe (schlechte Leistungen in allen Fächern) und eine "differential" Gruppe (gute Leistungen im Rechnen, schlechte Leistungen im Lesen). Es wurden zwei verschiedene AN-Verfahren eingesetzt, die als "Lesetest" (Auffinden von Synonymen mit multiple choice Antworten) und als "Rechentest" (Addition von jeweils 4 zweistelligen Zahlen) deklariert waren. Die einzeln durchgeführten Versuche waren auf drei Sitzungen im Abstand von je 2 - 3 Tagen verteilt. In einer ersten Sitzung erhielt die Vp unter "neutralen" Bedingungen (das heißt ohne manipulierte Leistungsrückmeldung) beide AN-Aufgaben. Die erreichte Leistung wurde in Sekunden/pro Durchgang angegeben. Die zweite Sitzung fand für je 6 Vpn aus den drei Versuchsgruppen unter "Erfolgsbedingungen" statt, während die anderen 6 Vpn jeder Versuchsgruppe unter "Mißerfolgsbedingungen" arbeiteten. Die "Erfolgs-" bzw. "Mißerfolgsbedingungen" wurden über eine fingierte, im voraus festgelegte Leistungsrückmeldung und über positive bzw. negative Leistungsbewertungen durch den V1 eingeführt. Der Vp wurde eine der beiden in Parallelformen vorliegenden AN-Aufgaben vorgelegt. Die dritte

Versuchssitzung verlief analog zur zweiten, nur die Bedingung ("Erfolg"/"Mißerfolg") und die dargebotene AN-Aufgabe (Parallellform) wechselten gegenüber der zweiten Sitzung.

Als Maß für das AN wird der Median der individuellen Zieldiskrepanzen - pro AN-Verfahren und pro Versuchsbedingung - bestimmt. Die Ergebnisse sind in Tabelle 1 zusammengefaßt.

Tabelle 1: Mittelwerte und Standardabweichungen der Zieldiskrepanzen der drei Versuchsgruppen (Erfolgsgruppe, Mißerfolgsgruppe, gemischte Gruppe) unter den Versuchsbedingungen "Neutral", "Erfolg", "Mißerfolg" (nach SEARS 1940)

AN-Verfahren	Versuchsbedingungen	Erfolgsgruppe		Mißerfolgsgruppe		gemischte Gruppe	
		N	$\bar{X}$ s	N	$\bar{X}$ s	N	$\bar{X}$ s
"Lesetest"	Neutral (Gesamtgruppe)	12	1,92 1,59	12	5,50 6,14	12	1,88 1,15
	Erfolg Neutral (1. Versuchssitzung)	6 ⁺	1,00 0,57	6	1,17 1,06	6	1,17 0,36
	Mißerfolg Neutral (1. Versuchssitzung)	6 ⁺	1,50 0,50	6	7,83 6,18	6	2,33 0,57
	Mißerfolg Neutral (1. Versuchssitzung)	6	2,83 2,61	6	3,75 3,17	6	4,17 2,42
"Rechentest"	Neutral (Gesamtgruppe)	12	2,17 2,46	12	6,71 8,56	12	4,54 5,04
	Erfolg Neutral (1. Versuchssitzung)	6	0,67 0,47	6	3,17 2,32	6	2,42 2,71
	Mißerfolg Neutral (1. Versuchssitzung)	6	2,17 3,28	6	6,08 8,98	6	1,42 4,24
	Mißerfolg Neutral (1. Versuchssitzung)	6	3,67 1,62	6	9,58 5,97	6	9,33 4,87
		6	2,17 1,17	6	7,33 8,08	6	7,67 3,62

+) Hier handelt es sich um dieselben Vpn. Die Ergebnisse unter Erfolgsbedingungen werden den zugehörigen Ergebnissen unter neutralen Bedingungen gegenübergestellt. Entsprechendes gilt für die übrigen Vpn-Gruppierungen, die sich aus der jeweiligen Kombination von AN-Verfahren und Versuchsbedingung ergeben.

Unter neutralen Bedingungen zeigen Vpn mit schlechten Schulleistungen (Mißerfolgsgruppe) gegenüber Vpn mit guten Schulleistungen (Erfolgsgruppe) im Durchschnitt höhere Zieldiskrepanzen (für den "Lesetest": $t = 2,18$; $P \leq .05$). Die Streuung der Zieldiskrepanzwerte differiert deutlich zwischen den drei Versuchsgruppen. Beim "Lesetest" sind die Unterschiede in allen Fällen statistisch gesichert. Für den Vergleich Erfolgs- versus Mißerfolgsgruppe, der hier vorrangig interessiert, heißt dies, daß die Zieldiskrepanzen bei den schlechten Schülern erheblich stärker streuen als bei den guten Schülern.

Für die Vergleiche zwischen den Versuchsbedingungen fallen die Resultate der Tendenz nach gleichsinnig, allerdings etwas weniger eindeutig aus.

SEARS hat in ihrer methodisch sehr umsichtig geplanten Untersuchung die Variablen Intelligenz und Schulleistung lediglich zur Vpn-Gruppierung herangezogen und bei der Datenauswertung nicht weiter berücksichtigt. Gezieltere Aussagen zur Beziehung der drei Variablen lassen sich hieraus also nicht ableiten. Überdies ist zu bedenken, daß SEARS mit recht kleinen Stichproben (12 bzw. 6 Vpn in den verschiedenen Untergruppen) gearbeitet hat.

Im Rahmen einer umfangreicheren Untersuchung (GOULD 1938, 1939), die der Beantwortung verschiedener Fragen galt, wurden an einer Stichprobe von College-Studenten ($N = 82$ männliche Vpn) Daten zum AN, zur Intelligenz und zur Studienleistung erhoben und korrelationsstatistisch analysiert (GOULD und KAPLAN 1940).

Das AN wurde mit sechs verschiedenen Aufgaben ("addition", "symbol-digit", "words", "steadiness", "cancellation"

und "target") untersucht, die nach den Angaben der Autoren rechnerische, mechanische, verbale und motorische Fähigkeiten messen. - Nähere Erläuterungen zu den AN-Aufgaben und -Versuchen sind bereits an anderer Stelle gegeben worden (siehe dazu S. 10). - Als AN-Maß wählte man die Zieldiskrepanz. Um ein einziges, für die sechs AN-Aufgaben repräsentatives Maß zu erhalten, wurden die individuellen Zieldiskrepanzwerte zunächst durch eine Transformation in Standardwerte ($\bar{X} = 50$; $s = 14$) vergleichbar gemacht. Für diese transformierten Werte wurde dann pro Vp der Median ermittelt. Die Autoren begründen diese Verrechnung zu einem kombinierten Maß mit der Höhe der Interkorrelationen der verschiedenen AN-Verfahren. (Die Koeffizienten liegen zwischen .04 und .44. Vgl. dazu auch S. 10.)

Zur Feststellung der Intelligenz wurden Testdaten herangezogen, die man bereits zu einem früheren Zeitpunkt (bei Eintritt in das College) mit Hilfe des "Thorndike Test of Scholastic Ability for High School Graduates" erhoben hatte. - Nähere Angaben zu diesem Intelligenztest, namentlich zu seinen psychometrischen Güteeigenschaften, werden nicht gemacht.

Als Indikator der Studienleistungen dienten vierstufig vergebene Leistungsbeurteilungen, die für jedes Fach vorlagen. Diese Benotungen wurden in drei Fächerkategorien (Naturwissenschaften und Mathematik, Sozialwissenschaften, Englisch und andere Sprachen) untergliedert und innerhalb jeder Kategorie gemittelt. Außerdem bestimmte man für jede Vp eine Durchschnittsnote über alle Fächer. (Es wird leider nicht mitgeteilt, auf welchen Zeitraum sich die Noten beziehen.)

Die errechneten Korrelationen werden in der folgenden Tabelle wiedergegeben.¹⁾

Tabelle 2: Korrelationen der Zieldiskrepanzwerte für die 6 AN-Verfahren mit den Gesamtdurchschnittsnoten und den Intelligenztestwerten (nach GOULD und KAPLAN 1940)

Anspruchsniveau-Aufgaben	Gesamtdurchschnittsnoten ^{+))}	Intelligenztestwerte (THORNDIKE)
Words	-.12 (82) ⁺⁺⁾	+++)
Steadiness	.07 (82)	
Addition	.00 (82)	
Symbol-digit	.02 (70)	
Cancellation	.12 (82)	
Target	-.17 (82)	
Median ZD-Wert	-.01 (82)	-.12 (78)

+) GOULD und KAPLAN begründen die Verwendung eines einzigen Durchschnittsmaßes mit hohen Interkorrelationen zwischen den drei Fächerkategorien (vgl. dazu S. 24).

++) Die Zahl in Klammern bezeichnet die der Berechnung jeweils zugrundeliegende Anzahl von Vpn. Diese Anzahl variiert, bedingt durch Stichprobenschwund bzw. fehlende Daten.

+++) Die AN-Aufgaben wurden nicht einzeln mit der Intelligenz in Beziehung gesetzt. Es wurde lediglich die Korrelation zwischen den Median-Zieldiskrepanzwerten und den Intelligenztestkennwerten gebildet.

Die durchweg recht niedrigen Korrelationen lassen keinerlei Regelhaftigkeit der Beziehungen erkennen. Aus diesen Ergebnissen klare Aussagen abzuleiten, ist nicht möglich, zumal auch die Angaben zur Datenerhebung und -auswertung

1) Zur Art der berechneten Korrelation und zur Signifikanz werden keine Angaben gemacht.

zum Teil unzureichend sind. Hinzu kommen einige am Zustandekommen der Ergebnisse möglicherweise beteiligte Bedingungen, deren Einfluß sich weder in seinem Ausmaß noch in seiner Richtung abschätzen läßt. (Zum Beispiel enthielt die Stichprobe ausschließlich männliche Vpn. Die Intelligenz- und die AN-Daten wurden mit einem relativ großen zeitlichen Abstand - etwa 1 bis 2 Jahre - erhoben, der außerdem zwischen den Vpn nicht übereinstimmte. - Dies ist aus folgendem zu schließen: Der Intelligenztest wurde bei Eintritt in das College durchgeführt. Bei den Studenten der untersuchten Stichprobe handelte es sich hauptsächlich um "sophomores", das heißt Studenten im 2. Studienjahr.) Schließlich besteht eine weitere, allerdings unumgängliche Schwierigkeit in bestimmten wahrscheinlichkeitstheoretischen Implikationen der Nullhypothese: "This is simply one of those fairly common scientific dilemmas where conclusions could be drawn from positive results but where nothing can be proved by negative results" (GOULD und KAPLAN 1940, S. 38).

In einer Untersuchung von WALTER und MARZOLF (1951), die in der Hauptsache der Frage nach Geschlechtsunterschieden hinsichtlich der AN-Setzung gilt, wird in einem Nebenergebnis mitgeteilt, daß "gute" und "schlechte" Schüler in ihrem durchschnittlichen AN (Zieldiskrepanz) nicht bedeutsam voneinander abweichen. Die Daten wurden an 80 Schülern des vierten, sechsten, achten und zwölften Schuljahres (pro Klassenstufe nach Zufall jeweils 10 Jungen und 10 Mädchen) erhoben. Zur Untersuchung des ANs verwendeten die Autoren das von ROTTER (1942) entwickelte "Aspiration Board", eine dem Billardspiel ähnliche Aufgabe. Als Kriterium der schulischen Leistung wählte man bereits vorliegende Ergebnisse aus Schulleistungstests (4. und 6. Klasse: "Metropolitan Achievement

Test"; 8. Klasse: "Stanford Achievement Test"; 12. Klasse: nicht näher bezeichnetes Verfahren aus einer Testbatterie). Für den Vergleich "guter" und "schlechter" Schüler wurden die Vpn innerhalb jeder Klassenstufe nach ihren Testwerten mediandichotomiert, da die Ergebnisse aus den verschiedenen Tests untereinander nicht vergleichbar waren.

WALTER und MARZOLF bemerken zu ihren Ergebnissen (S. 291): "The fact that level of scholastic achievement has so little effect upon goal discrepancy score, or level of aspiration, may come as somewhat of a surprise. However, these data indicate that academic achievement is not as potent an influence as we may have thought." Sieht man davon ab, daß es zumindest als fragwürdig gelten muß, die Schulleistung als unabhängige Variable zu betrachten, und bedenkt weiter, daß die Schulleistung nicht über Lehrerurteile, das heißt Noten, sondern über Tests erfaßt wurde, so erscheint diese recht allgemein gehaltene Folgerung der Autoren durchaus nicht zwingend. Denn es ist anzunehmen, daß die für den Schüler bzw. für sein Erleben ausschlaggebende schulische Realität sehr viel faßbarer in seinen Noten als in seinen Testergebnissen repräsentiert ist.

HELM (1967) unternahm eine Korrelationsstudie, in die auch die Variablen AN, Intelligenz und Studienleistung einbezogen waren. Zur Untersuchung des ANs verwendete HELM Pauli-Rechenbögen, die von den Vpn (N = 27; 12 männliche und 15 weibliche Studenten) unter Zeitbegrenzung zu bearbeiten waren. Als Intelligenzmaß wurde der Verbal-IQ des HAWIE (Hamburg-Wechsler-Intelligenztest für Erwachsene) ermittelt. Für die über Noten erfaßte Studienleistung wurde ein Globalwert (Notenmittel aller Hauptfächer) bestimmt.

HELM erhält für die Beziehung zwischen Intelligenz und Studienleistung ein r von .29 (n.s.). Die Korrelation des ANS ($|ZD|$, das heißt Verrechnung der Zieldiskrepanzen ohne Vorzeichen) mit der Studienleistung bleibt mit einem Wert von $r = -.37$ knapp unterhalb des 5%-Niveaus (r müßte .38 betragen, um bei 5 % signifikant zu sein). HELM sieht in der sich hierin andeutenden Tendenz eine Bestätigung der Ergebnisse von SEARS (1940). Demnach sind bei besseren Leistungen niedrigere Zieldiskrepanzen, bei schlechteren Leistungen eher hohe (positive oder negative) Zieldiskrepanzwerte zu finden (vgl. dazu S. 20-22).

Eine weitere deutsche Untersuchung, die am Konzept des Over- bzw. Underachievements orientiert ist, wurde in jüngerer Zeit von SCHERER und SCHLIEP (1974) vorgelegt. Die Autoren knüpfen an eine Untersuchung zur HAWIK-Standardisierung (SCHMALOHR 1962) an, in der bei der Hälfte der IQ-Werte von Volksschülern (IQ von 80 bis 100) und von Hilfsschülern (IQ von 77 bis über 100) eine Überschneidung festgestellt worden war (nach SCHERER und SCHLIEP 1974, S. 82), und stellen die Frage, "inwieweit sich jene auf der Sonderschule für Lernbehinderte befindenden 'underachiever' von Volksschülern unterscheiden" (SCHERER und SCHLIEP 1974, S. 82).

Zur Klärung dieser Frage war der Einfluß der Intelligenz konstant zu halten. Es wurden zwei hinsichtlich der Intelligenz parallele Gruppen der 5. Klasse einer Hauptschule ($N_1 = 22$) und einer Sonderschule für Lernbehinderte ($N_2 = 12$ (11)¹⁾) gebildet. Um möglichst homogene Stichproben zu erhalten, wurden ausschließlich

1) Bei der Darstellung der Ergebnisse wird durchgängig der auf $N = 11$ reduzierte Stichprobenumfang angegeben.

Jungen einbezogen. Die Intelligenz wurde mit dem "Primary Mental Ability Test" (PMA) von THURSTONE in der deutschen Bearbeitung von KEMMLER und LANGHEINRICH (1967) erfaßt. Die in Standardwerten ($\bar{X} = 100$; $s = 10$) anzugebende Testleistung wurde jeweils für den Gesamttest sowie für die sieben Untertests ermittelt. Die Ergebnisse der Parallelisierung, auf die wir noch genauer eingehen werden, sind in der folgenden Tabelle zusammengestellt.

Tabelle 3: Mittelwerte und Standardabweichungen der PMA-Werte für die Gruppe der Hauptschüler ($N_1 = 22$) und der Sonderschüler ($N_2 = 12$) (nach SCHERER und SCHLIEP 1974)

PMA	Hauptschüler		Sonderschüler		$\bar{X}_1 - \bar{X}_2$	$p^{+)}$
	$\bar{X}_1$	s_1	$\bar{X}_2$	s_2		
Gesamtstandardwert (Untertest 1-6) ⁺⁺⁾	89,72	4,16	89,91	7,42	-0,19	
Bilderwortschatz-Test	93,57	6,17	94,25	5,61	-0,68	
Wort-Test	94,00	6,57	84,00	6,75	10,00	<.001
Raumauffassungstest	93,18	8,39	95,66	10,26	-2,48	
Wortgruppen-Test	91,22	8,00	90,16	5,38	1,06	
Figurengruppen-Test	93,81	7,49	90,33	9,35	3,48	
Wahrnehmungsschnelligkeits-Test	89,86	8,63	100,91	13,04	-11,05	<.01
Zahlen-Test	93,50	10,76				

+) Für den Lokationsvergleich wurde der Mann-Whitney-U-Test verwendet.

++) Im Zahlentest zeigten 5 der 12 Sonderschüler extrem geringe Leistungen, so daß ihre Testwerte nicht zu ermitteln waren. Der Gesamtstandardwert basiert folglich nur auf sechs Untertests (ausgenommen Zahlentest).

Das AN wurde mit Hilfe von Additionsaufgaben (Addieren einstelliger Zahlen) im Gruppenverfahren untersucht. Die Vpn sollten innerhalb einer vorgegebenen Zeitbegrenzung möglichst viele Aufgaben bearbeiten. Zur Verrechnung der AN-Daten wählten die Autoren vier verschiedene Kennwerte: "Zieldifferenz" (mittlere Differenz zwischen den AN- und den Leistungswerten, relativiert auf die durchschnittliche Leistung), "Zieldiskrepanz" (im Unterschied zur Zieldifferenz werden hier die Vorzeichen der einzelnen Differenzen mit berücksichtigt), "starre Reaktionen" (Beibehalten eines ANs trotz Erfolg oder Mißerfolg¹⁾), "atypische Reaktionen" (Heraufsetzen des ANs nach Mißerfolg und Herabsetzen des ANs nach Erfolg¹⁾). In Tabelle 4 werden die zum AN erzielten Ergebnisse wiedergegeben.

Tabelle 4: Mittelwerte und Standardabweichungen der AN-Kennwerte für die Gruppe der Hauptschüler ($N_1 = 22$) und der Sonderschüler ($N_2 = 11$) (nach SCHERER und SCHLIEP 1974)

AN-Kennwerte	Hauptschüler		Sonderschüler		$\bar{X}_1 - \bar{X}_2$	p
	$\bar{X}_1$	s_1	$\bar{X}_2$	s_2		
Zieldifferenz	13,14	6,32	22,63	14,33	-9,40	<.05
Zieldiskrepanz	-3,28	9,19	3,36	21,97	-6,64	
Starre Reaktionen	1,61	1,79	0,90	0,90	0,71	
Atypische Reaktionen	0,42	0,73	1,54	1,31	1,12	<.001

1) Eine Definition für Erfolg und Mißerfolg wird nicht gegeben.

Gegenüber den Hauptschülern weisen die Sonderschüler signifikant größere Zieldifferenzen auf und zeigen häufiger atypische Reaktionen.

Vergegenwärtigt man sich die Ergebnisse der Parallelisierung (siehe Tabelle 3, S. 29), so ergeben sich Zweifel an der Aussagekraft der zum AN gefundenen Resultate. Parallelität der beiden Gruppen ist zwar in bezug auf den PMA-Gesamtwert - ausschließlich des Zahlentests - gegeben (wobei dies für die auffällig voneinander abweichenden Streuungen unklar bleibt), nicht aber in bezug auf die Untertests. Signifikante Unterschiede zeigen sich beim Wort-Test und beim Wahrnehmungsschnelligkeitstest. Im Zahlentest erreichten 5 Vpn der Sonderschulgruppe nicht einmal die für eine Testauswertung erforderliche Mindestleistung. Hierbei ist Überdies zu bedenken, daß die im Wort- und im Zahlentest abverlangten Leistungen für den Schulerfolg durchaus relevant sein dürften. Problematisch ist auch der sehr ungleiche Stichprobenumfang der Gruppen (22 Hauptschüler gegenüber 11 Sonderschülern). - Trotz der unzureichenden Parallelität beider Gruppen wird die Intelligenz von den Autoren als konstante und damit vernachlässigbare Größe betrachtet. Alle im weiteren gezogenen Schlußfolgerungen werden auf andere Variablen gegründet.

Auf Grund der dargestellten methodischen Probleme kann die von SCHERER und SCHLIEP gestellte Frage nicht als klar beantwortet gelten; wobei einzuräumen ist, daß die Forderung nach Parallelität sowohl im Gesamtwert wie in den Untertestwerten bereits die Grenzen des methodisch Machbaren überschreitet.

2.5. Entwicklung der Fragestellung

Fragt man nach den konstitutiven Bedingungen für das Zustandekommen geistiger, speziell schulischer Leistungen, so lassen der bisherige Verlauf und Ertrag der Forschung erkennen, daß hier ein erkundendes Anfangsstadium offensichtlich noch nicht überschritten worden ist. Dies gilt zum einen für die Untersuchung der motivationalen Faktoren der Schulleistung, darüber hinaus aber auch für die Beziehung der Variablen Motivation, Intelligenz und Schulleistung zueinander. Hierfür mögen neben inhaltlichen Gründen nicht zuletzt meßtechnische Probleme verantwortlich sein.

Im folgenden sollen die Grundgedanken dargestellt werden, die den Ansatz der vorliegenden Arbeit - inhaltlich und methodisch - fundieren.

Als eine allgemeine Tatsache war es wohl immer bekannt, daß die Schulleistung nicht nur von der Intelligenz, sondern auch von Motivationsfaktoren ("Fleiß") abhängt. Dennoch hat man die Motivation im Zusammenhang mit der Schulleistung bisher kaum zum Gegenstand von Untersuchungen gemacht (siehe dazu auch Punkt 2.4., S. 17 ff.). Vielmehr lag der Akzent der bis zum gegenwärtigen Stand geleisteten Forschung deutlich im Bereich der intellektuellen Voraussetzungen der Schulleistung.

Sucht man nach den Gründen für diese weitgehende Vernachlässigung der motivationalen Seite der Schulleistung, so ist - historisch betrachtet - zunächst auf den Einfluß des Behaviorismus und die dort vertretene Gegenstandsauffassung hinzuweisen. In der Konsequenz dieses auf das "objektiv" Faßbare verkürzten Forschungs-

ansatzes wurde die Motivation Überwiegend auf ihren physiologischen Aspekt reduziert (Vigilanz- und Aktivierungsforschung) oder zum Teil als Untersuchungsgegenstand überhaupt ausgeschlossen. Zwar hat man sich in der Tiefenpsychologie - bereits vor diesem behavioristischen Verdikt - in verstärktem Maß mit Fragen der Antriebe menschlichen Handelns auseinandergesetzt, jedoch waren hier klinische Gesichtspunkte vorrangig.

Eine allgemeine Psychologie der Motivation ist, verglichen mit der Wahrnehmungs-, Lern- und Denkforschung, erst relativ spät entstanden. Sieht man von der Willens- und Handlungspsychologie älteren Stils ab, so hat LEWIN damit begonnen, die Motivation zum Gegenstand der allgemeinen Psychologie zu machen. LEWIN hat die motivationalen Grundlagen des Handelns - sowohl in seinen experimentellen wie in seinen theoretischen Arbeiten - zu einem zentralen Thema gemacht. Die in den Jahren 1926 bis 1934 von ihm herausgegebenen "Untersuchungen zur Handlungs- und Affektpsychologie" sind hier als grundlegend zu nennen. Zugleich dokumentieren diese Untersuchungen, daß auch methodisch vergleichsweise schwierig zu handhabende Probleme, wie die der Motivation, einer experimentellen Behandlung zugänglich zu machen sind. LEWIN hat - in den von ihm angeregten Forschungen über das AN - geeignete methodische Grundlagen für die Untersuchung der leistungsbezogenen Motivationslage bereitgestellt.

Doch offensichtlich sind die im Konzept des ANs ursprünglich (z. B. von HOPPE, der die ersten Versuche zum AN vorlegte) erschlossenen methodischen Möglichkeiten zunehmend aus dem Blickfeld geraten, und zwar in dem Maß, als die AN-Forschung sich zur Leistungsmotivationsforschung hinentwickelt hat.

Auf der Suche nach tauglichen Meßverfahren zur Erfassung der Leistungsmotivation hat man neue Wege eingeschlagen. Im wesentlichen wurden zwei Gruppen von Verfahren entwickelt: die indirekten projektiven Verfahren des TAT und die direkt arbeitenden Fragebogenmethoden. Die erwartete Leistungsfähigkeit dieser Methoden ließ sich indessen weder bei den TAT-Verfahren noch bei den Fragebogenverfahren in einem befriedigenden Maß belegen. Beide Verfahrensgruppen weisen jeweils spezifische Mängel auf. (Zur ausführlicheren Darstellung und Kritik der TAT- und Fragebogenmethoden siehe auch Punkt 2.3., S. 12 ff.).

Die für den TAT-Ansatz grundlegende Annahme besagt, daß es dem erlebenden und handelnden Subjekt nicht möglich sei, direkte Mitteilungen über seine Leistungsmotivation zu machen, "da ... (es) ... sich von seinen eigenen Bezugssystemen nicht wie ein unabhängiger Beobachter distanzieren kann" (HECKHAUSEN 1963 b, S. 615). Hieraus wurde der Schluß gezogen, daß das indirekte projektive Verfahren als der bestmögliche und zugleich allein legitime Zugang zur Leistungsmotivation aufzufassen sei. - HECKHAUSEN (1963 a, S. 28 f.) illustriert diesen Gedankengang: "Um die zentralen Motive und Beweggründe eines Menschen möglichst rein und unvermischt zu erfassen, kommt alles darauf an, sie möglichst 'quellnahe anzuzapfen', ehe sie durch weitere Zuflüsse der verschiedensten Verhaltensdeterminanten trübe und undurchsichtiger werden."

Zunächst einmal wird deutlich, daß dieser Aspekt des Psychischen der Erfassung von Leistungsbereitschaft, Leistungsvertrauen durch die dem tiefenpsychologischen Ansatz verpflichtete Diagnose eine gravierende Wandlung

erfahren hat: dem Prinzip nach vergleichbar einer Wandlung von einer im Sinn von MÜNSTERBERG (oder auch von GIESE) durchgeführten Leistungsprobe, bei der die Leistungssituation in einen einfachen, überschaubaren Rahmen gebracht wird - experimentell zugänglich wird -, hin zu einem Reden über Leistung. Das heißt - um die Analogie weiterzuführen - Indikator der Leistung sind nicht mehr die aus der Leistungssituation heraus sich ergebenden Daten des einzelnen, sondern Aussagen über Inhalte, die in einem wie auch immer gearteten Zusammenhang mit Leistung stehen sollen - der Theorie nach. Die Gründe, die für das Beschreiten eines solchen Umwegs geltend gemacht werden, mögen durchaus triftig sein, bedauerlicherweise jedoch haben sie zu einer zunehmend größer werdenden Entfernung von der experimentellen Psychologie im engeren Sinn geführt. Erforscht wird nunmehr, bezogen auf den TAT-Ansatz, das theoretische Konstrukt "Leistungsmotivation", dessen phänomenale Repräsentation unklar ist. - Hierin ist wohl eine der Hauptursachen für die von TRAXEL geübte Kritik zu sehen, der in seinen Arbeiten die Emotionen als motivationale Erlebnisrepräsentanten in den Vordergrund stellt (TRAXEL 1961, 1962).¹⁾ - Die Durchsicht der Literatur ergibt keinen stringenten Beweis, daß es im Bereich der Emotionen, des Wollens und des Handelns eine separate Klasse von Erscheinungen gibt, die einen Begriff wie Leistungsmotivation auszufüllen in der Lage wäre.

Das wichtigste Argument, das für den TAT angeführt wird, sagt aus, daß man über die eigene Leistungsmotivation

1) Entsprechendes gilt wohl für die Kritik von DÜKER, der sein 1975 erschienenes Werk ja nicht ohne Grund "Untersuchungen über die Ausbildung des Wollens" genannt hat.

(das heißt implizit über Leistungsbereitschaft, Interesse, Engagement) keine direkten Urteile abgeben könne. Dies könnte zwei mögliche Ursachen haben (wobei, die erste a priori als richtig anzunehmen, durchaus nicht als zwingend zu betrachten ist): es wäre möglich, daß die die Leistungsmotivation konstituierenden Bedingungen dem Individuum tatsächlich nicht zugänglich sind - wie HECKHAUSEN behauptet und vermutlich auch experimentell geprüft hat; es könnte sehr wohl aber auch sein, daß die für die Gewinnung solcher Daten erforderlichen Skalen nicht ausreichend experimentell abzusichern sind. - Als Beispiele für eine solche Absicherung oder, präziser gesagt, für das Erfragen des "richtigen" Erlebnisrepräsentanten sind die Untersuchungen von TRAXEL (1960, 1961, 1962, 1966) und - aus dem sozialpsychologischen Bereich - von KRÜGER (1973, 1976) zu nennen.

Zieht man ein Fazit aus der soweit dargestellten und diskutierten Argumentation zum projektiven Ansatz, so muß die mit der Wahl des TAT gezogene methodische Konsequenz weder als die bestmögliche noch als die allein gangbare Lösung angesehen werden. Das Konzept des ANs - mit der Möglichkeit eines wohl eher phänomennahen Zugangs zur Leistungsmotivation - bietet sich als Alternative an. Der entscheidende Vorzug des AN-Konzepts liegt offensichtlich darin, daß im AN-Versuch tatsächlich eine Leistungssituation geschaffen wird, die sich zu Aussagen nutzen läßt über Bedingungen, die als konstituierend angesehen werden können für das, was man Leistungsmotivation nennt. Die im Versuch vorgegebene Situation gibt die Möglichkeit, die Dynamik des ANs, der aktuellen Motivationsvorgänge in einer für die Vp unmittelbar erlebbaren Leistungssituation in

Erscheinung treten zu lassen. Gegenüber den im TAT provozierten Phantasieproduktionen mit einer offenbar nur lockeren Realitätsverknüpfung¹⁾ wird die Vp im AN-Versuch dazu aufgefordert, konkrete Zielsetzungen vorzunehmen. In diesen Leistungsvornahmen, die jeweils unmittelbar vor der anschließend auszuführenden Leistung erfaßt werden, kommt ein als verpflichtend erlebter Anspruch an die eigene Tüchtigkeit zum Ausdruck (vorausgesetzt die Vp verhält sich instruktionsgemäß). Damit stehen die AN-Setzungen zugleich für das Maß an Anstrengung, das ein Individuum in einer gegebenen Leistungssituation aufzubieten bereit ist. Im wesentlichen ist es dieser Umstand der persönlichen Verbindlichkeit der Zielsetzungen und der daraus sich ergebenden engen Beziehung zum Handeln, zur Leistungssituation sowie die Möglichkeit des experimentellen Zugriffs und der quantitativen Behandlung der Daten, die für die Wahl des ANs sprechen. Zudem berechtigen die genannten Voraussetzungen wohl zu der Erwartung, daß sich das AN als ein hinreichend reliabler und valider Ausdruck der Leistungsmotivation (im Sinne tatsächlich wirksamer Ansprüche an die eigene Leistung) erweisen wird.

Faßt man den Stand der Forschung, an dem die vorliegende Arbeit ansetzt, zusammen, so ist folgendes festzustellen: Die Frage nach den motivationalen Anteilen der Schulleistung (im engeren Sinn) wurde erst sehr spät zum Gegenstand von Untersuchungen gemacht; und

1) In der TAT-Instruktion wird die Vp ausdrücklich dazu angehalten, sich beim Geschichtenerzählen etwas möglichst Phantasievolles einfallen zu lassen (vgl. HECKHAUSEN 1963 a, S. 58 f.).

die Schwierigkeiten ihrer methodischen Umsetzung haben dazu geführt, daß man bis dato einer befriedigenden Lösung nicht nähergekommen ist.

Es ist daher ein erstes Ziel dieser Arbeit, den motivationalen Gesichtspunkt der Schulleistung zur Geltung kommen zu lassen und zu untersuchen. Dazu mußte - als ein methodischer Schwerpunkt der Untersuchung - die Frage behandelt werden, wie diese motivationalen Bedingungen am angemessensten zu fassen seien. Das AN schien uns hier die Methode der Wahl zu sein.

Mit der Entscheidung für das AN stellt sich die Frage nach der intraindividuellen Konsistenz des ANs über verschiedene Leistungen (Generalität) und über die Zeit (Stabilität). Es geht hier um Grundlagenprobleme, die wohl, um ein sinnvolles Arbeiten mit dem AN-Konzept zu ermöglichen, nicht ausgeklammert werden können. In der vorliegenden Arbeit sollen diese Probleme, deren Klärung bislang nicht zufriedenstellend gelungen zu sein scheint (siehe auch Punkt 2.2., S. 6 ff.), daher weiter verfolgt werden.

Ausgehend von der Frage nach den konstituierenden Bedingungen (auf der Seite des Schülers) für das Zustandekommen der Schulleistungen, verfolgt diese Arbeit vorrangig das Ziel, die Beziehung zwischen den Variablen AN, Intelligenz und Schulleistung näher zu erkunden. - Die Problematik der Messung der Intelligenz und der Schulleistung wurde dabei (als bereits hinlänglich behandelt) zugunsten des ANs zurückgestellt. - Die wenigen zu der zu untersuchenden Beziehung vorliegenden Arbeiten (siehe auch Punkt 2.4.2., S. 18 ff.)

stehen recht unverbunden nebeneinander und sind zum Teil (in bezug auf die einbezogenen Variablen und/oder die methodische Umsetzung) nur bedingt einschlägig zu nennen. - Fast wäre zu fragen, ob es überhaupt einen Stand des Problems gibt. -

3. Methodik

Die vorliegende Untersuchung soll zur Klärung folgender Fragestellungen beitragen:

1. In welcher Beziehung stehen die Variablen AN, Intelligenz und Schulleistung?

Die Frage nach der intraindividuellen Konsistenz wird unter zwei Gesichtspunkten behandelt:

2. Ist das AN hinsichtlich verschiedener Leistungen relativ konsistent, oder variiert es je nach Art der zu erbringenden Leistung? (Generalität des ANs)
3. Wie konsistent (bzw. inkonsistent) ist das AN über die Zeit? (Stabilität des ANs)

Für die Untersuchung der in Frage stehenden Generalität des ANs war zu berücksichtigen, daß die zur Erfassung des ANs einzusetzenden Verfahren hinsichtlich der Art der jeweils abverlangten Leistung unterschiedlich sein sollten. Zur Schätzung der Stabilität des ANs wurde für die AN-Verfahren eine Versuchswiederholung vorgesehen. Hierbei sollte das Zeitintervall zwischen erster und wiederholter Versuchsdurchführung genügend groß gewählt werden, um etwaige Merkmalsschwankungen auch tatsächlich zur Wirkung gelangen zu lassen.

3.1. Entwicklung der Anspruchsniveauverfahren

Das AN dient uns als Indikator der individuellen Leistungsmotivation, und zwar im Sinne phänomenal gegebener, tatsächlich wirksamer Ansprüche an die eigene Leistung. - Die entscheidenden Argumente, die für die Wahl dieses methodischen Ansatzes geltend zu machen sind,

wurden bereits an anderer Stelle ausführlich dargestellt (siehe Punkt 2.5., S. 32 ff.) und sollen hier nicht nochmals aufgegriffen werden.

Zur Entwicklung der AN-Verfahren wurde eine Reihe von Kriterien¹⁾ herangezogen, die sich sowohl auf die Gestaltung der einzelnen Verfahren als auch auf deren Konstellation beziehen:

1. Die Vp muß vor tatsächliche Leistungsaufgaben gestellt werden, da es in einer Situation ohne Leistungscharakter nicht zur Bildung eines ANs kommt - wie sich schon in den HOPPEschen Versuchen (1930) zeigte.

Diese leistungsbetonte Situation läßt sich zum einen über bestimmte versuchstechnische Maßnahmen schaffen (z. B. Ausführen der Leistung unter Zeitbegrenzung), zum anderen - und vor allem - über die Instruktion induzieren.

2. Die zu stellende Aufgabe sollte an den Ehrgeiz oder allgemein an das Selbstbewußtsein der Vp appellieren (das heißt die Vp "in ihren Bann ziehen"), so daß es für die Vp von Bedeutung ist, wenn ihr im Versuch eine Leistung gelingt bzw. mißlingt.

Diesem Kriterium, das sich auf in der Vp herzustellende Bedingungen richtet (und damit nur indirekt das Aufgabenmaterial betrifft), Rechnung zu tragen, ist in erster Linie eine Frage der Instruktion.

1) Diese Kriterien haben sich teils aus der Kenntnis einschlägiger Untersuchungen zum AN, teils aus eigenen Versuchserfahrungen ergeben.

3. Es sollen mehrere, im Hinblick auf die jeweils abverlangte Leistung verschiedenartige AN-Aufgaben eingesetzt werden, um eventuelle Abhängigkeiten des ANs von der jeweiligen Art der zu erbringenden Leistung in Erscheinung treten zu lassen. - Die Berücksichtigung dieser Forderung zielt auf die Frage nach der Generalität des ANs.
4. Die Aufgaben müssen für die Vp eine angemessene mittlere Schwierigkeit besitzen, das heißt sie dürfen der Vp weder zu leicht noch zu schwer erscheinen. - Dies zu fordern ist insofern wichtig, als es nur innerhalb einer bestimmten Grenzzone der individuellen Leistungsfähigkeit zu AN-Bildungen kommt (siehe dazu auch unter Punkt 2.1. S. 5 unten) Nur in dieser Grenzzone der Leistungsfähigkeit liegen die Leistungsergebnisse, für die sich die Vp verantwortlich fühlt, die sie als Folge ihrer persönlichen Tüchtigkeit bzw. Untüchtigkeit erlebt.

Ein individuell mittlerer Schwierigkeitsbereich läßt sich im Einzelversuch direkt bestimmen und muß im Gruppenexperiment durch entsprechende Vorversuche bereichsweise angenähert werden.

5. Die unter Punkt 4. genannte Forderung impliziert, daß die Vp in der Lage sein muß, Anstieg und Abfall ihrer Leistung selber zu beurteilen.
6. Es sollen kurze Versuchsdurchgänge möglich sein, um eventuelle Störeinflüsse durch eine zu lange Versuchsdauer auszuschalten. Vor allem an zwei Fehlerquellen ist zu denken: Bei fortwährender Abfolge des gleichen Vorgangs ist eine Änderung der inneren Bedingungen zu erwarten, und zwar dergestalt, daß die Frage nach dem AN von der Vp mit einem zunehmenden

Grad an Bewußtheit aufgefaßt wird. Zudem ist mit Effekten der Ermüdung, der Langeweile, nachlassender Anspannung usw. zu rechnen.

7. Die Leistung sollte durch Übung in mehreren Durchgängen verbesserungsfähig (das heißt nicht zu sehr überlernt) sein; Leistungssteigerungen während des Versuchsablaufs sollten also möglich sein. Doch sollte dieser Übungsgewinn nicht stetig steigen (im Sinne eines "Man wird halt einfach besser!"), sondern immer wieder Schwankungen unterliegen, so daß erwartet werden kann, daß Erfolgs- und Mißerfolgsergebnisse häufiger wechseln und die Vp in einer gewissen Spannung gehalten wird.
8. Noch wichtiger - gegenüber der Verbesserungsfähigkeit durch Übung - ist wohl die Möglichkeit, durch eigene willentliche Anstrengung eine Leistungssteigerung zu erreichen. Gerade die Haltung des Sichselbst-aktivieren-Könnens, das Erlebnis "Es hängt von mir ab", scheint für einen AN-Versuch ideal zu sein.
9. Zu den bereits angesprochenen Schwankungen im Verlauf der Leistungskurve (siehe oben unter 7.) ist - bezogen auf den eher kognitiven Aspekt des ANs - folgendes zu ergänzen: Die Variabilität der Leistungen von Durchgang zu Durchgang sollte nicht zu groß sein, um eine annähernde Vorhersage der künftigen Leistung zu ermöglichen; andererseits sollte der Zusammenhang von einem zum nächsten Durchgang nicht so eng sein, daß die nächsten Leistungswerte ganz sicher vorausgesagt werden können. - Beides, ein Zuviel ebenso wie ein Zuwenig an Variabilität, würde einer echten AN-Bildung entgegenwirken.

Ausgehend von diesen Überlegungen wurde eine Reihe von verschiedenen AN-Aufgaben in Betracht gezogen und erprobt. Es kamen schließlich drei Verfahren zur Auswahl:

1. Zahlenaddition (paarweises Addieren einstelliger Zahlen),
2. Zahl-Symbol-Substitution (Zuordnen von Symbolen zu Zahlen nach einem bestimmten Schlüssel),
3. Ringewerfen (Ringe über einen Stab werfen).

Gemeinsam ist allen drei Verfahren, daß es sich um relativ einfache, wiederholt durchführbare Aufgaben handelt, die sich in ihrer Schwierigkeit innerhalb desselben Verfahrens nicht voneinander unterscheiden. Das Spektrum der von der Aufgabe her möglichen Leistungen (Bereich zwischen minimal und maximal möglicher Leistung) wurde jeweils so bemessen, daß die interindividuelle und die intraindividuelle Variation der Leistungen sich uneingeschränkt darstellen konnte. (Hierzu wurden Vpn herangezogen, deren Altersstufe der für die Hauptversuche geplant entsprach.) Das AN wird als ausdrückliche Zielsetzung für eine gleichbleibende Aufgabentätigkeit erhoben. Die Vp muß sich für eine bestimmte Höhe des zu erreichenden Zieles bei einer erneuten Ausführung der gleichen Aufgabe entscheiden. Die vorgenommene Leistung kann der anschließend erreichten entsprechen oder diese um einen bestimmten Betrag (Anzahl der Aufgaben) über- bzw. unterschreiten. Anspruchs- und Leistungsniveau werden in Einheiten derselben Leistungsskala (Anspruchsniveau: angestrebte Anzahl zu bearbeitender Aufgaben; Leistungsniveau: tatsächlich erreichte Anzahl zu bearbeitender Aufgaben) für jeden Durchgang erfaßt.

Die Formulierung der Frage nach dem AN wirft gewisse Probleme auf. Sie kann, wenn sie ungerichtet formuliert ist, von den Vpn recht verschieden aufgefaßt werden. Die Art der Auffassung kann sich bewegen zwischen dem, 'was man zu schaffen hofft', und dem, 'was man (realistischerweise) zu schaffen erwartet'.¹⁾ In dieser Weise konfundierte Auffassungsweisen würden die Interpretierbarkeit der so gewonnenen Daten sehr fragwürdig werden lassen. (Die Probleme wären ähnlich denen, wie sie sich im Fall des projektiven Tests ergeben.) Es mußte daher eine für die Vp eindeutiger zu verstehende Formulierung gefunden werden. Mit der schließlich von uns gewählten Frage "Was schätzt du, ...?" sollte eine realistischere Haltung der Vp gegenüber dem eigenen Leistungsanspruch induziert werden. In bezug auf das hier anstehende Problem, konkret Stellung zur eigenen Leistungsanforderung zu beziehen, schien uns die ausgesuchte Frage die einschlägigere, die "natürlichere" zu sein - gegenüber der Frage nach dem eigenen Leistungswunsch, die eher unverbindlich und daher als weniger geeignet anzusehen ist. Das Einschätzen der zukünftigen eigenen Leistung stellt sich dar als etwas, was durch das nachfolgende Eintreten der Leistung bestätigt oder nicht bestätigt wird, wohingegen das Wünschen einer Leistung nicht eine solche Rückkopplung über die Leistung erfährt (das heißt nicht mit dem konkreten Leistungshandeln korrespondiert).

1) Untersucht wurde diese Frage der Auffassung z. B. von FESTINGER (1942). Für Vpn, die gefragt wurden "What score would you like to get next time?" ergab sich im Mittel eine signifikant größere Zieldiskrepanz als für Vpn, die gefragt wurden "What score do you expect to get next time?". Außerdem zeigte die 'like'-Gruppe in ihren Zieldiskrepanzwerten eine bedeutend höhere Streuung als die 'expect'-Gruppe.

Im Hinblick auf die Aussagefähigkeit der Ergebnisse erschien es uns wünschenswert, eine Kontrolle darüber zu haben, ob es sich bei den von der Vp gemachten Angaben um echte AN-Setzungen und nicht etwa nur um irgendwelche beliebigen Mitteilungen handelt. Eine bloße Mitteilung, ohne daß eine innerlich übernommene Zielsetzung vorläge, kann dort zur Annahme eines ANs führen, wo in Wirklichkeit keines vorhanden ist. Aus der Tatsache, daß das Auftreten von Erfolgs- und Mißerfolgs-erlebnissen darüber Auskunft gibt, ob tatsächlich ein AN gebildet wurde¹⁾, ergibt es sich, die "Echtheit" der Zielsetzungen über das jeweils zu erfassende Erleben von Erfolg und Mißerfolg zu kontrollieren. Um also das AN nicht zu "operational" zu fassen, soll dieses Urteil über erlebten Erfolg bzw. Mißerfolg - im Sinne eines Gültigkeitskriteriums - als ein kritischer Indikator für die phänomenal gegebene Situation fungieren.

Es stellt sich die Frage, wie eine solche Urteilsabgabe zweckmäßigerweise gestaltet sein sollte. Nachdem in den bisher zum AN vorgelegten Untersuchungen auf eine derartige Kontrolle fast durchweg verzichtet wurde²⁾, kann der innerhalb der vorliegenden Arbeit unternommene Versuch, eine Gültigkeitskontrolle einzuführen, nicht mehr als nur ein erster, erkundender und

1) Diese Beziehung zwischen dem AN und dem Erleben von Erfolg und Mißerfolg wurde bereits in der ersten Untersuchung zum AN festgestellt (HOPPE 1930) und später als Ergebnis immer wieder repliziert (vgl. dazu auch S. 5).

2) Es fanden sich lediglich zwei Arbeiten, in denen ein Urteil über Erfolgs- bzw. Mißerfolgserleben erhoben wurde, allerdings im einen Fall als Fremdurteil des Vls (JUCKNAT 1937), im anderen Fall in einer Untersuchung über "einige emotionale, conative und zeitliche Bedingungen des Erinnerns" (JÄGER 1958).

daher vorläufiger Beitrag zur Lösung dieses Problems sein. - Die Hauptschwierigkeit besteht darin, die in Rede stehenden Erlebnisdaten möglichst unverfälscht zu erfassen. Besonders im Versuch mit Kindern - die vorliegende Untersuchung soll an Zehn- bis Zwölfjährigen (Schüler der 5. Klasse) durchgeführt werden (siehe dazu Punkt 3.2.1., S. 54 ff.) - ist es wichtig, Urteilkategorien zu finden, die bei den Kindern eine erlebnismäßige Repräsentanz haben und zudem geeignet sind, eine Steigerung bzw. Abstufung des Erlebens tatsächlich darzustellen. Die Begriffe Erfolg und Mißerfolg direkt zu verwenden, erscheint insofern nicht ratsam, als sie bereits zu sehr mit bestimmten, die Ergebnisse verzerrenden Wertungen verknüpft sind. Es fällt wohl nicht schwer, das Erlebnis eines Erfolges mitzuteilen, dagegen kostet es vielleicht einige Überwindung, einen erlebten Mißerfolg, der ja ohnehin schon eine Belastung des Selbstwertgefühls bedeutet, dann auch noch einzugestehen. (Entsprechende Beobachtungen konnten bei Zehn- bis Zwölfjährigen und, in einer anderen Untersuchung, auch schon bei Vier- bis Sechsjährigen gemacht werden.) Im Hinblick auf diese Gefahr, statt eines Erlebens die Bewertung eines Erlebens zu erfassen, kommt hinzu, daß im Zusammenhang mit Leistung diese für sich genommen schon als ein Wert betrachtet wird - was einen möglichst unmittelbaren Zugang zum Erleben in bezug auf diese Leistung nicht eben erleichtert.

Beziehen wir die skizzierten Überlegungen auf die anstehende Frage, so muß die der Vp vorzugebende sprachliche Bezeichnung, Umschreibung so "passend" wie möglich das einfangen, was der Vp erlebbar ist (nicht mehr und nicht weniger), wobei dies ebenso naheliegend wie

schwierig erscheint. Für die konkrete Umsetzung bedeutete dies, daß die Vp am Ende eines jeden Durchgangs (das heißt nach einer AN-Setzung mit anschließend ausgeführter Leistung) angeben mußte, wie zufrieden sie war mit dem, was sie gerade geschafft hatte. Diese Formulierung hatte sich als die am ehesten geeignete erwiesen, nachdem zuvor versuchsweise verschiedene andere Formulierungen verwendet worden waren. In Vorversuchen an einer ausgewählten Stichprobe wurde das Urteil zunächst zweistufig ("zufrieden"/"nicht zufrieden") erprobt (vgl. dazu unter Punkt 3.2. S. 53 f.). Die dort gemachten Beobachtungen legten nahe, eine weitere Abstufung des Urteils einzuführen, um eine hinreichende Differenzierung zu ermöglichen. Es wurde daher in den Hauptversuchen eine vierstufige, bipolare Skala ("sehr zufrieden"/"zufrieden"/"unzufrieden"/"sehr unzufrieden") mit einem ausgegliederten Punkt ("weiß nicht") vorgegeben (vgl. dazu unter Punkt 3.3.3.1. S. 66).

Es wurden innerhalb der einzelnen AN-Verfahren jeweils elf Durchgänge durchgeführt. Diese Anzahl hatte sich unter folgenden Gesichtspunkten als zweckmäßig herausgestellt: Es sollte einerseits eine in Hinsicht auf die Zuverlässigkeit der Messung genügend breite Datenbasis, vor allem aber sollten die Voraussetzungen dafür geschaffen werden, die Dynamik des ANs in ihrem Verlauf faßbar zu machen; andererseits sollte - wie dies bei einer sehr großen Anzahl von Durchgängen der Fall wäre - vermieden werden, den Vpn eine zu lange Versuchsdauer zuzumuten (vgl. dazu auch das unter Punkt 6. Gesagte auf S. 42 f.).

Während die eingesetzten AN-Verfahren in den bisher beschriebenen Merkmalen (formaler Ablauf des Versuchs, Frage nach dem AN, Urteil über erlebten Erfolg/Mißerfolg, Anzahl der Durchgänge pro Verfahren) untereinander übereinstimmen, unterscheiden sie sich in bezug auf die Art der Aufgabe bzw. der zu erbringenden Leistung (vgl. dazu Punkt 3. auf S. 42).

Beim "Zahlenaddieren" geht es um einfache arithmetische Rechenoperationen, die, sieht man von der schreibmotorischen Komponente ab, in erster Linie Konzentration bei einer recht gleichförmigen Tätigkeit verlangen. Die Vp muß (in Anlehnung an den Pauli-Test¹⁾) je zwei einstellige Zahlen, die in senkrechten Reihen angeordnet sind, miteinander addieren. Die Zahlen von 1 bis 9 sind innerhalb jeder Reihe in eine Zufallsabfolge gebracht, wobei eine direkte Aufeinanderfolge gleicher Zahlen ausgeschlossen war, um den Schwierigkeitsgrad der Additionsaufgaben untereinander annähernd gleich zu gestalten. Eine Reihe entspricht einem Durchgang und umfaßt 30 Zahlen. Es kommt für die Vp darauf an, innerhalb einer pro Durchgang vorgegebenen Zeit (die über alle Durchgänge des Versuchs konstant gehalten wird) möglichst viele Additionsaufgaben zu bearbeiten. Über die Festsetzung der Zeitbegrenzung sollte zweierlei bewerkstelligt werden: 1. Es sollte einerseits ein Deckeneffekt, andererseits ein Sohleneffekt vermieden werden, was nicht nur auf die tatsächlich erreichten Leistungen, sondern auch auf die Zielsetzungen zu beziehen ist (siehe dazu auch Punkt 4. auf S. 42).

1) Als neuere Literatur zu diesem Test sei auf das von W. ARNOLD verfaßte Handbuch (1970 erschienen als vierte verbesserte und erweiterte Auflage von R. PAULI: Der Arbeitsversuch als charakterologisches Prüfverfahren. Leipzig 1943.) hingewiesen.

Dies bedeutet, daß der qua Aufgabe vorgegebene Schwierigkeitsbereich - HOPPE (1930, S. 25) nennt dies "Aufgabenniveau" - dem einzelnen einen genügend weiten Spielraum lassen muß für seine individuell zu findende mittlere Schwierigkeitszone. 2. Die günstigsten Voraussetzungen für eine hohe Zuverlässigkeit sind dann gegeben, wenn die Varianz ein Maximum annehmen kann. Um dies zu ermöglichen, sollte die durchschnittlich erreichte Leistung der Vpn auf halber Höhe der maximal möglichen Leistung liegen. Eine unter diesen Gesichtspunkten geeignete Zeitbegrenzung wurde in entsprechenden Vorversuchen ermittelt (siehe auch unter Punkt 3.2. S. 54 sowie unter Punkt 3.2.3.1. S. 65 und Punkt 3.2.3.2. S. 78).

Im Verfahren "Zahl-Symbol-Substitution" wird von der Vp vor allem assoziative Beweglichkeit, Umstellungsfähigkeit gefordert. Darüber hinaus spielen Konzentration, kurzfristiges Merken und visuell-motorische Koordination eine Rolle. Wie beim "Zahlenaddieren" ist auch hier ein schreibmotorischer Faktor enthalten. Die Aufgabe der Vp besteht darin, bestimmte Symbole bestimmten Zahlen nach einem vorgegebenen Schlüssel zuzuordnen. Insgesamt acht verschiedene Symbole (zusammengestellt nach dem Zahlen-Symbol-Test aus dem Hamburg-Wechsler-Intelligenztest für Erwachsene, BONDY 1964) sind den Zahlen von 1 bis 8 zugeordnet:

1	2	3	4	5	6	7	8
=	X	^	○	□	└	┐	—

Ein Durchgang setzt sich aus 24, in einer waagerechten Reihe angeordneten Zuordnungsaufgaben zusammen, so daß

jede der 8 Zahl-Symbol-Kombinationen dreimal vorkommt. Die Folge der Zuordnungen innerhalb eines Durchgangs ist zufällig - mit der Einschränkung, daß keine direkte Abfolge gleicher Zahlen auftreten sollte (es gilt hierfür die gleiche Begründung, wie bereits beim "Zahlenaddieren" dargestellt; siehe dazu S. 49). Bei einer pro Durchgang begrenzten Zeit geht es für die Vp darum, möglichst viele Symbole den vorgegebenen Zahlen zuzuordnen. Auch hier folgt die Festsetzung der Zeitbegrenzung denjenigen Gesichtspunkten, die bereits beim "Zahlenaddieren" genannt sind (siehe dazu S. 49 f.).

Ausschlaggebend für die beim "Ringewerfen" zu erbringende Leistung ist die motorische Geschicklichkeit und damit auch die visuell-motorische Koordination. Außerdem ist die Konzentration der Vp von Bedeutung. Es ist die Aufgabe der Vp, Ringe aus einer bestimmten Entfernung über einen Stab zu werfen. Die verwendeten Ringe haben einen Innendurchmesser von 9 cm, die Höhe des Stabes, der zur Fixierung in eine Platte eingelassen ist, beträgt 58 cm. Der Vp werden pro Durchgang jeweils 16 Ringe zur Verfügung gestellt, von denen sie dann soviel wie möglich über den Stab werfen soll. Die Leistungsmenge ergibt sich aus der Anzahl der Treffer. Eine Zeitbegrenzung wird nicht vorgenommen. Nachdem die Leistungen beim "Ringewerfen" eine beträchtliche interindividuelle Streuung aufweisen (wie sich in entsprechenden Versuchen mit zehn- bis zwölfjährigen Schülern gezeigt hat), wird der Bereich einer mittleren Schwierigkeit individuell ermittelt. In einem Probedurchgang, der dem eigentlichen Versuch vorausgeht, wirft die Vp alle 16 Ringe. Je nachdem, wie hoch die Leistung der Vp ausfällt, wird die Entfernung Vp - Stab so korrigiert,

daß der damit vorgegebene Schwierigkeitsbereich den Leistungen der Vp angemessen ist und in etwa eine mittlere Trefferquote erreicht werden kann. Sofern es sich im weiteren Verlauf des Versuchs als nötig herausstellt (das heißt sich die Aufgabe als "zu leicht" bzw. "zu schwer" erweist), wird eine nochmalige Korrektur der Entfernung vorgenommen.

Das Ringewerfen wurde im Einzelversuch, das Zahlenaddieren und das Zahl-Symbol-Substituieren konnten im Gruppenversuch durchgeführt werden. - Die Situation des Gruppenversuchs legt es möglicherweise für die Vp nahe, sich mit dem Platznachbarn über vorgenommene und erreichte Leistungen zu verständigen, was eine in ihrem Ausmaß nicht abzuschätzende Verzerrung der Ergebnisse zur Folge haben würde. Um diese Kommunikation zu unterbinden, wurden für die Durchführung der Gruppenversuche entsprechende Vorkehrungen getroffen (siehe dazu unter Punkt 3.2.3.1. S.63f). Damit schien uns ein in bezug auf das AN wesentlicher Unterschied zwischen der Situation des Gruppenversuchs und der des Einzelversuchs aufgefangen zu sein. Jedoch wird man davon ausgehen müssen, daß weitere Unterschiede zwischen Gruppen- und Einzelversuch bestehen. Es bleibt immer noch das Problem, daß die Leistungsvornahme einmal quasi öffentlich gesagt, im anderen Fall mehr privat notiert wird. - Die Möglichkeit, aus Gründen der Vergleichbarkeit im Einzelversuch ebenfalls die jeweils vorgenommenen und die erreichten Leistungen von der Vp notieren zu lassen, wurde zwar erwogen, letztlich aber als für die Einzelversuchssituation zu unnatürlich verworfen. - Diese Frage sich ergebender Unterschiede müßte im Grunde Gegenstand einer separaten Untersuchung sein.

3.2. Hauptversuche

Es wurden zunächst ausführliche Vorversuche an einer ausgelesenen Stichprobe unternommen. Vorrangiges Ziel dieser Vorversuche war die Erprobung der entwickelten AN-Verfahren. Es wurde eine Stichprobe von 20 Vpn (8 Mädchen, 12 Jungen) untersucht, die hinsichtlich ihrer Intelligenz- (Frankfurter Analogietest 4-6, BELSER, ANGER, BARGMANN und RAATZ 1973) und Recht-schreibleistung¹⁾ (Diagnostischer Rechtschreibtest 4-5, MEIS 1970) extreme Merkmalsausprägungen aufwiesen, womit sich folgende doppelte Extremgruppenbildung ergab:

		FAT	
		-	+
DRT	-	N=5	N=5
	+	N=5	N=5

Die Extremgruppenbildung sollte dazu dienen, etwaige Unterschiede zwischen den Untergruppen im Hinblick auf das AN eher in Erscheinung treten zu lassen. Nach einer ersten Versuchsdurchführung mit allen drei AN-Verfahren fand - mit einem zeitlichen Abstand von 14 Wochen - eine Versuchswiederholung statt. - Die Ergebnisse der Vorversuche führten zu einer Reihe von Veränderungen

-
- 1) Die als Gruppierungsmerkmal herangezogene Recht-schreibleistung steht hier als Indikator für die Schulleistung; dies erscheint aus folgenden Gründen statthaft: 1. In der zur Untersuchung herangezogenen Altersstufe (Schüler zu Beginn des 6. Schuljahres) nimmt die Rechtschreibleistung - durch ihre starke Gewichtung - noch eine Schlüsselstellung in bezug auf die Schulleistung ein. 2. DRT-Leistungen und schulische Diktatleistungen zeigen einen relativ hohen Zusammenhang.

bei den AN-Verfahren, die z. B. die Instruktion, die Zeitbegrenzungen für das Zahlenaddieren und das Zahl-Symbol-Substituieren sowie das Urteil über erlebten Erfolg/Mißerfolg betrafen.

3.2.1. Auswahl der Stichprobe

Die Stichprobe für die Hauptversuche sollte sich in bezug auf die Variablen Intelligenz und Schulleistung nicht nur (wie in den kurz geschilderten Vorversuchen) aus Vpn mit extremen Merkmalsausprägungen zusammensetzen; vielmehr sollte die gesamte Variationsbreite vorkommender Ausprägungsgrade einbezogen werden, um die zu untersuchenden Beziehungen auch im Bereich mittlerer Ausprägungsgrade erfassen zu können und damit eine größere Repräsentativität herzustellen. Günstige Voraussetzungen dafür, einen möglichst umfassenden Variationsbereich in die Untersuchung aufnehmen zu können, bieten sich in der Gesamtschule, wo das Leistungsspektrum aller drei traditionellen Schultypen repräsentiert ist.

Die Untersuchungen wurden in der Gesamtschule Hollfeld (Oberfranken) durchgeführt.¹⁾ Hinsichtlich der Klassenstufe entschieden wir uns für den 5. Schülerjahrgang. Die in dieser Schule eingerichteten sieben 5. Klassen

1) Dem Leiter der Schule, Herrn Oberstudiendirektor O. Hofmann, der der Arbeit sehr aufgeschlossen gegenüberstand, gilt mein besonderer Dank, ebenso dem stellvertretenden Leiter, Herrn A. Saam, ohne dessen Hilfe die organisatorisch recht aufwendige Eingliederung der Untersuchungen in den komplizierten Schulbetrieb einer Gesamtschule kaum möglich gewesen wäre.

mit insgesamt 207 Kindern umfassen alle Schüler des 5. Jahrgangs im Schulsprengel Hollfeld-Wonsees. Dies hat im Hinblick auf die gewünschte Leistungsbreite den Vorteil, daß eine durch den Übertritt in weiterführende Schulen bedingte Auslese noch nicht stattgefunden hat. Für die vorliegende Untersuchung ist weiter zu bedenken, daß das Gesamtschulsystem allgemein eine deutlich ausgeprägte Leistungsorientierung zeigt, was offensichtlich auch für die hier gewählte Schule zutrifft. Dies wird für die Interpretation der Ergebnisse sicherlich von Belang sein.

Es wurden sämtliche 5. Klassen in die Untersuchung einbezogen. Die Ausgangsschülerzahl reduzierte sich um 22 Vpn. 20 Schüler haben bei mindestens einem der durchgeführten Verfahren gefehlt. Zwei Schüler mußten ausgeschlossen werden, da sie die Instruktion (beim Zahlenaddieren bzw. beim Zahl-Symbol-Substituieren) falsch verstanden hatten und folglich ihre Daten nicht auswertbar waren. Somit standen für die Auswertung die Daten von 185 Vpn (101 Mädchen, 84 Jungen) zur Verfügung. Die Altersstruktur (Stichtag: Durchführung des ersten Verfahrens innerhalb der Gesamtuntersuchung) entspricht mit 11 Jahren $\pm$ 1 Jahr der Norm (ausgenommen eine Vp, die bereits 13 Jahre alt war). Die soziologische Struktur der Schüler erweist sich als relativ homogen. Entsprechend dem ländlichen Einzugsgebiet der Schule sind die Kinder überwiegend bäuerlicher Herkunft.

3.2.2. Erfassung der Intelligenz und der Schulleistung

Für die Messung der Intelligenz erschien es uns besonders wichtig, einen Test auszuwählen, bei dem die Testleistung möglichst wenig an dynamische Faktoren gebun-

den ist, um nicht bereits durch die Wahl des Intelligenztests einen Zusammenhang zum AN zu begünstigen. Dies läßt sich wohl durch die Wahl eines Tests beeinflussen, der eher auf Qualität (Niveau) als auf Quantität (Schnelligkeit) angelegt ist. Denn die Leistung in einem Niveautest hängt in weit geringerem Maß von motivationalen Bedingungen ab, als dies bei einem Schnelligkeitstest der Fall ist (vgl. LIENERT 1967). Außerdem sollte der einzusetzende Test, wenn möglich, nicht so konstruiert sein, daß er vorrangig bzw. ausschließlich eine "schulspezifische Intelligenz" mißt, womit die erst zu untersuchende Beziehung zwischen Intelligenz und Schulleistung in ihrem Ergebnis schon weitgehend vorweggenommen wäre. Dieser Gesichtspunkt läßt sich allerdings nur begrenzt berücksichtigen, da Intelligenztests in aller Regel auch an Schulnoten validiert werden. Ein weiteres Auswahlkriterium bestand in einer hinreichenden und möglichst in allen Leistungsbereichen gleich guten Differenzierungsfähigkeit des Tests. Schließlich war es im Sinne der Ökonomie wünschenswert, den Test als Gruppenverfahren durchführen zu können.

Aufgrund dieser Überlegungen wurde zur Intelligenzmessung der Frankfurter Analogietest 4-6 (BELSER, ANGER, BARGMANN und RAATZ 1973) herangezogen. Der FAT 4-6 erfaßt im wesentlichen den THURSTONEschen Faktor R (Reasoning), das heißt die Fähigkeit, Gesetzmäßigkeiten zu erkennen und anzuwenden. Die Analogien werden mit Wörtern, Bildern und Figuren gebildet. Die Autoren legen besonderen Wert auf die Erfahrungsunabhängigkeit der gestellten Aufgaben. Unabhängig vom Milieu und von der besuchten Schule sollen jedem Kind die verbalen Inhalte des Tests glei-

chermaßen verständlich sein. Es wurden spezielle Wortschatzuntersuchungen durchgeführt, um sicherzustellen, daß sich das Testergebnis primär auf Relationen stiftendes Denken bezieht und vom mehr oder minder umfangreichen Wortschatz der Kinder weitgehend unbeeinflusst ist. Sinngemäß - so die Autoren - gilt das gleiche für die beiden anderen im Test verwendeten Aufgabenarten (Bilder und Figuren). Der FAT 4-6 setzt sich aus 95 nach dem Schwierigkeitsgrad ansteigend geordneten Analogieaufgaben zusammen, die zu unterscheiden sind in Wort-, Bild-, Figuren- und Mischanalogien (letztere als Kombinationen der ersteren). Das Prinzip im Aufbau der Aufgaben besteht durchweg darin, daß ein vollständiges Paar zusammen mit einem unvollständigen Paar vorgegeben ist (z. B. Sommer - Regen : Winter - ...), wobei die Vp das fehlende Wort (bzw. Bild, Figur) des zweiten Paares aus jeweils fünf zur Auswahl stehenden Antworten heraussuchen muß, um die Analogie zwischen den beiden Paaren herzustellen. Der Test liegt in zwei Parallelförmigkeiten (A und B) vor und ist als Gruppenverfahren durchzuführen. Eine Zeitbegrenzung wird lediglich global vorgenommen; das heißt die zur Bearbeitung der Testaufgaben insgesamt zur Verfügung stehende Zeit ist auf 30 Minuten beschränkt, wobei die Vpn laut Instruktion über diese Zeitbegrenzung nicht informiert werden sollen. Der FAT 4-6 ist für Schüler des 4. Schuljahres an Grundschulen sowie für Schüler des 5. und 6. Schuljahres an Hauptschulen, Realschulen und Gymnasien bestimmt. Als Maß für die Intelligenz liefert der Test einen Gesamtwert, dessen Rohwert der Anzahl richtig gelöster Aufgaben entspricht. Normen liegen als T-Werte und Prozentränge vor für die verschiedenen Schuljahre und Schultypen. Die zu den Gütekriterien des Tests vorgelegten Ergebnisse sollen hier nur zusammenfassend

dargestellt werden (für detailliertere Angaben sei auf das Beiheft zum FAT 4-6 verwiesen; BELSER und RAATZ 1973). Hinsichtlich der Objektivität (in ihren verschiedenen Aspekten) genügt der Test den zu stellenden Anforderungen. Die Zuverlässigkeitsbestimmung nach KUDER-RICHARDSON ergab für alle Schüler (5. und 6. Klasse; Form A und B) einen Koeffizienten von $r_{tt} = .95$ (N=756 je Testform). Für die gleiche Stichprobe erbrachte eine Reliabilitätskontrolle nach dem Paralleltestverfahren (Zeitintervall: 1 Tag; Verwendung beider Testformen) einen Wert von $r = .91$. - Nach LIENERT (1967) sollten standardisierte Tests eine Konsistenz von $r_{tt} \geq .90$ und eine Paralleltestreliabilität von $r_{tt} \geq .80$ aufweisen. - Die Validität wurde an verschiedenen Außenkriterien überprüft. Interessant in bezug auf die vorliegende Arbeit sind die ermittelten Korrelationen zwischen den Leistungen im FAT 4-6 und Schulnoten (erfaßt als letzte Zeugnissnoten in den Fächern Deutsch und Mathematik).

Tabelle 5: Korrelationen (Produkt-Moment-Koeffizienten) zwischen den Testwerten im FAT 4-6 und den Zeugnissnoten in den Fächern Deutsch und Mathematik (nach BELSER und RAATZ 1973)

		Deutschnote		Mathematiknote	
		N	r	N	r
Haupt- schule	5. Klasse	450	0,34	548	0,40
	6. Klasse	383	0,38	414	0,44
Real- schule	5. Klasse	360	0,12	416	0,19
	6. Klasse	507	0,12	499	0,23
Gymna- sium	5. Klasse	899	0,27	913	0,25
	6. Klasse	831	0,23	832	0,29

Die Variable Schulleistung kann - der grundsätzlichen Möglichkeit nach - zum einen über Zensuren, zum anderen über Schulleistungstests erfaßt werden. Für die vorliegende Arbeit entschieden wir uns, Noten als Indikator der Schulleistung zu erheben. - Noten sind als Urteile des Lehrers mit all den sich daraus ergebenden, möglichen Fehlerquellen belastet. Die unter dem Gesichtspunkt der Gütekriterien, insbesondere von INGENKAMP vorgebrachte Kritik (zuletzt in einem Sammelband: INGENKAMP 1976) richtet sich vor allem auf eine mangelnde Objektivität und Zuverlässigkeit der Noten. Über die hieraus zu ziehenden Konsequenzen besteht allerdings nach wie vor kein Konsens. Somit ist die an der Notengebung geübte Kritik zwar durchaus nicht gegenstandslos, jedoch ist die Verwendung standardisierter Schulleistungstests als Alternative in anderer Hinsicht keineswegs weniger problematisch. Während die Testverfahren ein höheres Maß an Objektivität und Reliabilität erwarten lassen, ist hingegen zu bezweifeln, ob damit auch die Gültigkeit, als entscheidendes Gütekriterium, wesentlich verbessert werden kann. Das für die gewählte Verwendung von Noten letztlich ausschlaggebende Argument ist darin zu sehen, daß die für den Schüler, präziser gesagt für das Erleben des Schülers relevante schulische Realität wohl sehr viel faßbarer in seinen Noten als in seinen Testwerten repräsentiert ist; dies gilt namentlich dann, wenn - wie dies bei der vorliegenden Untersuchung der Fall wäre - die Testdaten lediglich für schulexterne Untersuchungszwecke erhoben werden und daraus keinerlei Konsequenzen für das schulische "Schicksal" des einzelnen entstehen. Anders ausgedrückt kann die Beziehung zwischen AN und Schulleistung wohl nur dann psychologisch sinnvoll untersucht werden, wenn für die

Variable Schulleistung ein Indikator gefunden wird, der die Schulleistung als das faßt, was für den Schüler wirklich ist, das heißt seine schulische Leistungserfahrung begründet.

Die Schulleistung wurde über die Zensuren im Zwischen- und im Jahreszeugnis des 5. Schuljahres erfaßt.¹⁾ Im Sinne des zuvor dargestellten Arguments sind Zeugnisnoten offenbar am ehesten geeignet, den Ernstcharakter der Schulleistung einzufangen. Außerdem haben Zeugnisnoten den Vorteil, als eine Art Durchschnittswert aus mehreren Noten stabiler und aussagekräftiger zu sein als eine Einzelnote. Die Noten wurden für alle, insgesamt elf Fächer erhoben:

- Religion
- Deutsch
- Englisch
- Mathematik
- Erdkunde
- Physik/Chemie²⁾
- Biologie
- Musik
- Kunsterziehung
- Handarbeit bzw. Werken
- Sport

Damit liegen als Ausgangsdaten von jeder Vp zweimal 11 Zeugnisnoten vor. In der Gesamtschule Hollfeld wird ab der zweiten Hälfte der 5. Jahrgangsstufe eine äußere Differenzierung in den Fächern Englisch und Mathematik, eine Binnendifferenzierung im Fach Deutsch durchgeführt. Daher sind bei den drei Differenzierungsfächern

1) Die Untersuchung wurde am Ende des 5. Schuljahres durchgeführt, so daß die entsprechenden Zeugnisnoten zur Verfügung standen.

2) Der Unterricht sieht Inhalte aus beiden Fächern vor.

im Jahreszeugnis zusätzlich zu den Noten die jeweiligen Kurseinstufungen (A, B, C oder SB¹⁾) aufgenommen.

Die zunächst in Betracht gezogene Möglichkeit, für jeden Schüler - als ein weiteres Datum zur Schulleistung - ein globales Lehrerurteil über die gesamte Breite schulischer Leistung zu erheben, erwies sich als nicht realisierbar. Bedingt durch die Organisationsform der Schule verfügte weder der Klassenlehrer noch sonst ein in der jeweiligen Klasse unterrichtender Lehrer über eine hinreichend umfassende Erfahrungsbasis, wie sie für die Abgabe eines solchen Urteils erforderlich gewesen wäre. Darüber hinaus erschien es fraglich, ob und inwieweit es überhaupt möglich ist, ein derartiges Schätzurteil unabhängig von den jeweiligen Noten der Schüler (die ja auch ein Lehrerurteil darstellen) abzugeben.

3.2.3. Durchführung der Versuche

Die Hauptversuche untergliedern sich in zwei Abschnitte: Im ersten Abschnitt wurden der FAT 4-6 sowie die drei AN-Verfahren durchgeführt (Hauptversuche I); im zweiten Abschnitt fand - zur Schätzung der Stabilität des ANs - nach 3 1/2 Monaten eine Versuchswiederholung der AN-Verfahren statt (Hauptversuche II).

Die Versuche wurden sämtlich in den Räumen der Gesamtschule, also einer für die Schüler vertrauten Umgebung, durchgeführt und so gut wie möglich in den normalen Schulbetrieb integriert.

1) SB: Schüler mit Sonderbetreuung.

3.2.3.1. Hauptversuche I

Die Versuche wurden am Ende des 5. Schuljahres im Juli 1976 (in den beiden ersten Wochen) ausgeführt. Soweit die Organisation des Unterrichts dies zuließ, wurden die Versuche in die mittleren Schulstunden gelegt, eine Zeit, wo nicht mehr bzw. noch nicht mit Ermüdung der Schüler zu rechnen ist. - Zwar nicht aus der Perspektive des wissenden Untersuchers, aber eben aus der Sicht der Vp (und das ist hier entscheidend) geht es nicht nur im FAT, sondern auch in den AN-Versuchen um - möglichst gut bzw. schnell zu erbringende - Leistungen. - Pro Versuchstag wurde jeweils nur ein Verfahren angesetzt, wobei die Abfolge so geregelt war, daß zuerst der FAT, dann alternierend die beiden AN-Gruppenverfahren (Zahlenaddition¹⁾, Zahl-Symbol-Substitution¹⁾) und als letztes der AN-Einzelversuch (Ringewerfen¹⁾) durchgeführt wurden. Bei den Verfahren FAT, ZA und ZS wurden die Schüler klassenweise zur Untersuchung herangezogen. Insgesamt waren sieben Klassen zu untersuchen. Um die Versuchszeit für die Verfahren über alle Klassen konstant halten und zudem sequentielle Vl-Effekte ausschalten zu können, wurde die Zuordnung der Vl (7)²⁾ zu den Klassen und den Verfahren systematisch variiert, wie im folgenden Schema dargestellt:

-
- 1) Im weiteren werden folgende Abkürzungen verwendet:
Zahlenaddition - ZA; Zahl-Symbol-Substitution - ZS;
Ringewerfen - RW.
 - 2) Die Versuche wurden von der Verfasserin zusammen mit zwei Studenten und vier Studentinnen durchgeführt. - Den Studenten, die mit beachtlichem Engagement bei der Sache waren, sei ganz herzlich gedankt.

Klassen Ver- fahren	1	2	3	4	5	6	7
FAT 4-6	I ⁺)	II	III	IV	V	VI	VII
ZA	II	III	IV	V	VI	VII	I
ZS	III	IV	V	VI	VII	I	II
RW	IV	V	VI	VII	I	II	III

+) V1 I bis VII

Im Interesse möglichst objektiver Durchführungsbedingungen wurde auf eine gründliche Vorbereitung der V1 besonderer Wert gelegt. Zu diesen Vorbereitungen gehörte zum Beispiel, daß der jeweilige Versuchsablauf, soweit er nicht bereits durch die Instruktion festgelegt ist, in allen Einzelheiten (einschließlich eventueller Vorkommnisse) besprochen wurde. Außerdem wurde das V1-Verhalten nach Art des Rollenspiels eingeübt. Für den Vortag des eigentlichen Versuchsbeginns war eine "Generalprobe" angesetzt, bei der alle Verfahren (FAT 4-6, ZA, ZS, RW) in einer 6. Klasse (Gesamtschule Hollfeld) durchgeführt wurden.

Insbesondere für die AN-Gruppenverfahren erschien es uns sehr wichtig, ein mögliches Abschreiben bzw. eine Verständigung der Schüler untereinander zu verhindern (zur Begründung siehe S. 52). Dazu wurden für die Durchführung der Gruppenverfahren einige Vorkehrungen getroffen. Zusätzlich zum V1 nahm pro Klasse ein Lehrer teil (jeweils derjenige Lehrer, der in der entsprechenden Schulstunde in der Klasse Unterricht gehalten hätte), der den V1 beim Aufpassen unterstützte. Die - möglichst weiträumige - Sitzordnung der Vpn

wurde anhand einer alphabetischen Klassenliste geregelt. Miteinander befreundete und daher untereinander eventuell "kommunikationsfreudigere" Schüler konnten damit nur noch durch Zufall nebeneinander sitzen.

Für die Durchführung des FAT 4-6 erübrigen sich weitere Erläuterungen, da der Ablauf durch die vorgegebene Instruktion bereits genau festgelegt ist (vgl. BELSER und RAATZ 1973, S. 4 ff.).

Die Versuchsanweisungen zu den drei AN-Verfahren, die in einer schriftlichen Fassung vorlagen, wurden vom V1 mündlich gegeben. Die mündliche Darbietung schien uns eher geeignet, die Vp anzusprechen, sie - im Sinne der Instruktion - für den Versuch zu gewinnen. Der jeweilige Versuchsablauf soll im folgenden anhand der Instruktionen näher beschrieben werden. (Je ein Exemplar des für die AN-Verfahren verwendeten Versuchsmaterials befindet sich im Anhang auf den Seiten 1 ff.)

Zum Zahlenaddieren:

"Bei dem, was wir gleich machen, wollen wir sehen, wie gut Ihr durchhaltet und wie schnell Ihr seid. Dazu bekommt Ihr einige Rechenaufgaben, die Ihr bestimmt alle könnt.

Bei den nun folgenden Aufgaben sollt Ihr die untereinanderstehenden Zahlen so zusammenzählen, wie ich es Euch jetzt an der Tafel vorführe. (V1 demonstriert anhand der Beispielreihe an der Tafel die jeweils paarweise Addition der senkrechten Zahlenreihen: 1. + 2. Zahl, 2. + 3. Zahl usw.; gemeinsam mit den Schülern rechnen.)

Ihr sollt also immer zwei untereinanderstehende Zahlen zusammenzählen, und das Ergebnis schreibt Ihr rechts zwischen die beiden Zahlen, so wie ich es hier an der Tafel gemacht habe. Es wird immer nur eine Zahlenreihe gerechnet. Nach jeder Reihe machen wir eine kurze Pause.

Versucht bitte, vor jeder Zahlenreihe zu schätzen, wie viele Aufgaben Ihr bei der nächsten Reihe schaffen könnt. Schreibt dann diese Zahl in das Kästchen oben über die Reihe, die Ihr anschließend rechnet (V1 zeigt auf oberen Kasten an der Tafel).

Ihr habt für jede Reihe 35 sec (also etwas mehr als eine halbe Minute) Zeit."

Ausgehend von der in den Vorversuchen (vgl. dazu S. 53 f.) in der ersten Versuchsdurchführung vorgegebenen Zeitbegrenzung und der dabei durchschnittlich von den Vpn erreichten Leistung, wurde für die Hauptuntersuchungen eine Zeitbegrenzung ermittelt, bei der im Durchschnitt eine Leistung erwartet werden konnte, die auf halber Höhe der maximal möglichen Leistung lag (zu den Gesichtspunkten für die Festsetzung der Zeitbegrenzung siehe S. 49 f.).

"Versucht, so schnell wie möglich, aber natürlich auch immer richtig zu rechnen."

Dieser Hinweis, daß es auf Schnelligkeit und Richtigkeit ankomme, schien uns geboten, um einer Konfundierung von Quantität und Qualität in bezug auf die vorgenommenen und erreichten Leistungen entgegenzuwirken. Die Vp sollte sich nur auf die jeweilige Anzahl zu lösender bzw. gelöster Aufgaben beziehen; der Aspekt der Richtigkeit sollte dabei als eine Selbstverständlich-

keit eingehen, das heißt keinen eigenen Bezugspunkt bilden.

"Ihr sollt immer der Reihe nach in 35 sec möglichst viele Aufgaben von einer Reihe rechnen. Wenn ich 'Achtung - los!' sage, fangt Ihr an zu rechnen. Wenn ich nach 35 sec 'Stop!' sage, hört bitte sofort auf zu rechnen und macht unter der zuletzt gerechneten Aufgabe einen Strich. Es darf dann nichts mehr verbessert werden. Nach dem 'Stop'-Zeichen machen wir eine kurze Pause. Dann zählt Ihr, wie viele Aufgaben Ihr gerechnet habt, und ganz unten in das Kästchen unter der Reihe (V1 zeigt auf unteren Kasten an der Tafel) schreibt Ihr, wie viele Aufgaben es waren.

Anschließend sollt Ihr angeben, wie zufrieden Ihr mit dem seid, was Ihr gerade geschafft habt. - Wenn Ihr z. B. mit dem, was Ihr gerade geschafft habt, 'sehr zufrieden' seid, dann macht Ihr 2 Pluszeichen (V1 zeigt auf dafür vorgesehenen Kreis an der Tafel), so wie Ihr es auf dem kleinen Blatt¹⁾ seht, das Ihr alle bekommen habt; wenn Ihr 'zufrieden' seid, macht Ihr 1 Pluszeichen; wenn Ihr 'unzufrieden' seid, macht Ihr 1 Minuszeichen; wenn Ihr 'sehr unzufrieden' seid, macht Ihr 2 Minuszeichen, und wenn Ihr mal nicht sagen könnt, ob Ihr zufrieden oder unzufrieden seid, dann macht Ihr eine Null.

Wenn Ihr dies gemacht habt, sollt Ihr schätzen, wie viele Aufgaben Ihr wohl bei der nächsten Reihe schaffen könnt. Schreibt diese Zahl wieder in das Kästchen oben über der nächsten Reihe und wartet dann, bis ich wieder 'Achtung - los!' sage.

1) Siehe Anhang S. 3.

So sollt Ihr es bei allen Reihen machen.

Habt Ihr das alles verstanden? (Eventuelle Fragen beantworten.)

(Eventuelle Wiederholung der Gesamtinstruktion in Stichworten.)

Vor dem 1. Durchgang

Wie Ihr die Aufgaben lösen sollt, das wißt Ihr jetzt. Man kann es schnell oder langsam machen. Ihr sollt so schnell wie möglich rechnen, aber dabei auch keine Fehler machen.

Klappt jetzt das Blatt, auf das Ihr Euren Namen geschrieben habt, zur Seite.

Wir fangen links mit der 1. Reihe an. Schätzt bitte zuerst, wie viele Aufgaben Ihr in der 1. Reihe wohl schaffen könnt und schreibt diese Zahl in das Kästchen über der 1. Reihe. Jede Reihe hat 29 Aufgaben. Bleistift hinlegen. (Warten bis alle fertig sind.) Jetzt fangen wir mit dem Rechnen an, Bleistift in die Hand. 'Achtung - los!'

Nach jedem weiteren Durchgang

'Stop!'. Strich machen. Gerechnete Aufgaben zählen und die Zahl unten in das Kästchen unter die Reihe schreiben. Angeben, wie zufrieden Ihr mit dem seid, was Ihr gerade geschafft habt. Schätzen, wie viele Aufgaben Ihr bei der nächsten Reihe wohl schaffen könnt und die Zahl in das Kästchen über die nächste Reihe schreiben. Wenn Ihr dies gemacht habt, Bleistift hinlegen, damit ich sehen kann, ob Ihr alle soweit seid. - Nun zur nächsten Reihe. Bleistift in die Hand. 'Achtung - los!'

Nach dem 11. (letzten) Durchgang

'Stop!'. Strich machen. Gerechnete Aufgaben zählen und die Zahl unten in das Kästchen unter die Reihe schreiben. Angeben, wie zufrieden Ihr mit dem seid, was Ihr gerade geschafft habt."

Die Versuchsdauer beim Zahlenaddieren betrug etwa 30 bis 40 Minuten.

Zum Zahl-Symbol-Substituieren:

"Heute machen wir eine andere Aufgabe als gestern.¹⁾ Dabei wollen wir sehen, wie gut Ihr Euch etwas merken könnt und wie schnell Ihr Euch auf etwas Neues umstellt.

Es sind Aufgaben, die Ihr sicher alle könnt.

Auf dem Blatt, auf das Ihr eben Euren Namen geschrieben habt, seht Ihr eine Reihe mit abgeteilten Kästchen (an der Tafel auf den Zuordnungsschlüssel zeigen). Im oberen Kästchen steht immer eine Zahl, und im unteren Kästchen steht immer ein Zeichen. Es sind die Zahlen von 1 bis 8. Zu jeder Zahl gehört ein ganz bestimmtes Zeichen. (V1 nennt die einzelnen Zuordnungen und zeigt sie an der Tafel.)

Darunter seht Ihr eine Kästchenreihe, in der nur Zahlen stehen; die dazugehörenden Zeichen fehlen noch. Ihr

1) Das Zahlenaddieren und das Zahl-Symbol-Substituieren wurden an aufeinanderfolgenden Tagen durchgeführt.

sollt jetzt immer der Reihe nach¹⁾ zu den Zahlen die dazugehörigen Zeichen eintragen. (V1 erfragt zu den einzelnen Zahlen die Zeichen und trägt diese auf der Tafel ein.)

Dies war eine Beispielaufgabe zum Üben. Bei den folgenden Aufgaben, mit denen wir gleich beginnen, sollt Ihr es dann genauso machen. Immer der Reihe nach von links nach rechts tragt Ihr zu der Zahl das dazugehörige Zeichen ein. Es kommen in allen Reihen nur die Zahlen mit den Zeichen vor, wie Ihr sie schon aus unserer Beispielaufgabe kennt. Auf jedem Blatt ist immer nur eine Reihe, in die Ihr die fehlenden Zeichen einsetzen sollt. Außerdem findet Ihr auf jedem Blatt wieder die Reihe mit den Zahlen von 1 bis 8 und den dazugehörenden Zeichen.²⁾

Versucht bitte, vor jeder Reihe zu schätzen, wie viele Kästchen Ihr wohl bei der nächsten Reihe ausfüllen könnt. Schreibt dann diese Zahl in den Kasten links neben die Reihe, die Ihr anschließend bearbeitet. (V1 zeigt auf linken Kasten an der Tafel.)

Ihr habt für jede Reihe 18 sec Zeit (also etwas mehr als eine Viertelminute). Versucht, die Zeichen so schnell wie möglich, aber natürlich auch immer richtig einzutragen."

-
- 1) In Vorversuchen hatten einige Vpn statt die Aufgaben in der vorgegebenen Zufallsabfolge zu lösen, zuerst alle 1er-Zuordnungen, dann alle 2er-Zuordnungen usf. bearbeitet, was zwar nicht den offensichtlich gewünschten Effekt besserer Leistungen hatte, aber in der Konsequenz die Vergleichbarkeit gegenüber den anderen Vpn problematisch machte.
 - 2) Mit dieser Maßnahme, den Zuordnungsschlüssel in jedem Durchgang (das heißt auf jedem Blatt) erneut darzubieten, sollte eine zwischen den Durchgängen vergleichbare Schwierigkeit erreicht werden.

Die Zeitbegrenzung wurde analog dem beim ZA beschriebenen Vorgehen ermittelt (siehe dazu S. 65). Als Begründung für den Hinweis auf die anzustrebende Schnelligkeit und Richtigkeit der Leistung gilt hier Entsprechendes wie bereits beim ZA angeführt (siehe dazu S. 65 f.).

"Ihr sollt in 18 sec möglichst viele Kästchen von einer Reihe ausfüllen. Wenn ich 'Achtung - los!' sage, fangt Ihr an (links in der Reihe, 1. Kästchen). Wenn ich nach 18 sec 'Stop' sage, hört bitte sofort auf und macht hinter dem zuletzt eingetragenen Zeichen einen langen senkrechten Strich. Es darf dann nichts mehr verbessert werden. Nach dem 'Stop'-Zeichen machen wir eine kurze Pause. Dann zählt Ihr, wie viele Zeichen Ihr eingetragen habt, und in den Kasten rechts neben die Reihe schreibt Ihr, wie viele Zeichen es waren (V1 zeigt auf rechten Kasten an der Tafel)."

Der Teil der Instruktion, der sich auf das Urteil über erlebten Erfolg bzw. Mißerfolg bezieht, hat den gleichen Wortlaut wie beim ZA (siehe dazu S. 66).

"Wenn Ihr dies gemacht habt, sollt Ihr schätzen, wie viele Zeichen Ihr wohl bei der nächsten Reihe schaffen könnt. Schreibt diese Zahl wieder in den Kasten links neben die Reihe auf das nächste Blatt und wartet dann, bis ich wieder 'Achtung - los!' sage.

So sollt Ihr es dann bei allen Reihen machen.

Habt Ihr das alles verstanden? (Eventuelle Fragen beantworten.)

(Eventuelle Wiederholung der Gesamtinstruktion in Stichworten.)

Vor dem 1. Durchgang

Wie Ihr die Aufgaben lösen sollt, das wißt Ihr jetzt. Man kann es schnell oder langsam machen. Ihr sollt es so schnell wie möglich, aber auch ohne Fehler machen.

Klappt jetzt das Blatt, auf das Ihr Euren Namen geschrieben habt, zur Seite und schätzt bitte zuerst, wie viele Zeichen Ihr in der 1. Reihe wohl schaffen könnt. - Es sind in jeder Reihe 24 Zeichen. - Schreibt die Zahl, die Ihr Euch vornehmt, in den Kasten links neben die Reihe. Bleistift hinlegen. (Warten bis alle fertig sind.) Jetzt fangen wir an, Bleistift in die Hand, 'Achtung - los!'

Nach jedem weiteren Durchgang

'Stop'. Senkrechten Strich machen. Eingetragene Zeichen zählen und die Zahl in den Kasten rechts neben die Reihe schreiben. Angeben, wie zufrieden Ihr mit dem seid, was Ihr gerade geschafft habt. Schätzen, wie viele Zeichen Ihr bei der nächsten Reihe wohl schaffen könnt und die Zahl in den Kasten links vor die nächste Reihe schreiben. Wenn Ihr dies gemacht habt, Bleistift hinlegen, damit ich sehen kann, ob Ihr alle soweit seid. Nun zur nächsten Reihe. Bleistift in die Hand. 'Achtung - los!'

Nach dem 11. (letzten) Durchgang

'Stop'. Senkrechten Strich machen. Eingetragene Zeichen zählen und die Zahl rechts in den Kasten neben die Reihe schreiben. Angeben, wie zufrieden Ihr mit dem seid, was Ihr gerade geschafft habt."

Der Versuch dauerte durchschnittlich 20 bis 30 Minuten.

Zum Ringewerfen:

Die Einzelversuche wurden zur gleichen Zeit in sieben verschiedenen Versuchsräumen (die uns für die Dauer der Versuche großzügigerweise von der Schule zur Verfügung gestellt worden waren) durchgeführt. Die Vpn-Abfolge richtete sich nach einer alphabetischen Klassenliste. Mit dieser willkürlichen Regelung sollte - soweit möglich - verhindert werden, daß jeweils befreundete Schüler, die vermutlich eher geneigt sind, sich über den Versuch zu verständigen, nacheinander zum Versuch kommen. Der Vp wurde folgende Instruktion gegeben:

"Bei den Aufgaben, die wir heute machen, will ich einmal sehen, wie geschickt Du bist. Dazu sollst Du aus einer bestimmten Entfernung Ringe über diesen Stab werfen. (Vl zeigt auf Stab und Ringe.) Es sind insgesamt 16 Ringe. Versuch' bitte, soviel Ringe wie möglich über den Stab zu werfen. Du sollst die Ringe mehrere Male über den Stab werfen. Jedes Mal hast Du wieder 16 Ringe zur Verfügung, von denen Du soviel wie möglich über den Stab werfen sollst.

Vor jeder Wurfreihe sollst Du schätzen, wieviel von den 16 Ringen Du wohl über den Stab werfen kannst. Diese Zahl sagst Du mir dann jedes Mal.

Jedes Mal, wenn Du alle 16 Ringe geworfen hast, zählen wir, wieviel Du über den Stab gebracht hast.

Dann sollst Du mir bitte sagen, wie zufrieden Du mit dem bist, was Du gerade geschafft hast. Dabei sag' mir dann bitte, ob Du mit dem, was Du gerade geschafft hast,

'sehr zufrieden' oder 'zufrieden' oder 'unzufrieden' oder 'sehr unzufrieden' bist; und wenn Du mal nicht sagen kannst, ob Du zufrieden oder unzufrieden bist, dann sagst Du 'weiß nicht'. Du zeigst mir dann jedes Mal auf diesem Blatt (Blatt für Erfolg/Mißerfolg-Urteil) wie zufrieden Du bist."

Die in dieser Form gestaltete Urteilsabgabe (anhand des dazu vorgegebenen Blattes) sollte die Vp veranlassen, die vorgegebene Differenzierungsbreite des Urteils auch tatsächlich auszuschöpfen. Die optische Darbietung schien uns geeignet, bei der Vp alle Urteilkategorien gleichermaßen präsent zu halten. Zudem war diese Art der Urteilsabgabe eher derjenigen vergleichbar, die bei den Verfahren ZA und ZS gewählt worden war.

"So sollst Du es dann bei allen Wurfreihen machen. Hast Du das alles verstanden?

(Eventuelle Fragen beantworten.)

Vor dem O. Durchgang (Probedurchgang)

Zuerst müssen wir noch die Entfernung zum Stab ausmessen. Damit es für Dich nicht zu schwer und auch nicht zu leicht wird, soll die Entfernung nicht zu groß und auch nicht zu klein sein. (Entfernung Vp - Stab messen, markieren und notieren.) Jetzt machen wir erst mal eine Probereihe zum Üben, und dann fangen wir richtig an. (Vp wirft aus bestimmter Entfernung 16 Ringe. Je nach Anzahl der Treffer gegebenenfalls Distanz korrigieren.)¹⁾

1) Die dieser Maßnahme zugrunde liegenden Überlegungen wurden bereits an anderer Stelle dargestellt (vgl. dazu S. 42 Punkt 4. und S. 51 f.).

Vor dem 1. Durchgang

Jetzt geht es richtig los. Denk' dran, Du sollst soviel Ringe wie möglich über den Stab werfen. Vorher sag' mir aber bitte, wieviel Ringe Du wohl bei der 1. Wurfreihe über den Stab werfen kannst.

Nach jedem weiteren Durchgang

(LN¹) notieren.) Zeig' mir bitte, wie zufrieden Du bist mit dem, was Du gerade geschafft hast.

(E/M-Urteil²) notieren.) Was schätzt Du, wieviel Ringe Du wohl bei der nächsten Reihe schaffen kannst.

(AN-Angabe notieren.)

Nach dem 11. (letzten) Durchgang

(LN und E/M-Urteil notieren.) So, ... (Name) ..., Du bist jetzt fertig. Wenn Du jetzt in Deine Klasse zurückgehst, verrätst Du den anderen aber bitte noch nichts von dem, was wir gemacht haben. (Sehr wichtig ... Versprichst Du mir das? ... Sonst Ergebnisse unbrauchbar ...)."

Nachdem insbesondere bei Einzelversuchen, die sich über einen gewissen Zeitraum erstrecken (die Durchführung beanspruchte 4 Tage), es nicht auszuschließen ist, daß die Vpn sich über das Versuchsgeschehen verständigen und möglicherweise eine gegenseitige Beeinflussung stattfindet, hielten wir es für ratsam, diesen Still-

1) LN: erreichte Leistung (Anzahl der Treffer).

2) E/M-Urteil: Urteil über erlebten Erfolg bzw. Mißerfolg.

schweigeappell am Ende des Versuchs an die Vp zu richten. Die Vp sollte quasi als ein Verbündeter gewonnen werden, der um seine Verantwortung weiß und sich an seine Zusicherung zu schweigen gebunden fühlt. - Dieser Appell wurde offenbar weitgehend befolgt.

Die V1 sollten der Vp gegenüber ruhig, freundlich und sachlich auftreten und sich während des Versuchs jeglicher in ihrer Wirkung unkontrollierbarer Kommentare (z. B. Loben, besonderes Anspornen etc.) enthalten. Zusätzlich zu den quantitativen Versuchsdaten wurden freie Verhaltensbeobachtungen und spontane Äußerungen der Vp registriert, sofern diese in irgendeinem Zusammenhang mit der Versuchsaufgabe standen. Die Dauer eines Einzelversuchs betrug durchschnittlich 15 Minuten. Es wurde darauf geachtet, daß während der regulären Pausen keine Versuche stattfanden.

3.2.3.2. Hauptversuche II

Zur Kontrolle der Stabilität des ANs wurden die AN-Verfahren nach einem Zeitintervall von 3 1/2 Monaten wiederholt durchgeführt. Die Zeitspanne wurde so lang bemessen, um Erinnerungseinflüsse (die zu einer Überschätzung der Stabilität führen würden) möglichst auszuschalten und auf die Frage nach der Merkmalskonstanz eine befriedigende Antwort geben zu können (vgl. auch unter Punkt 3. S.40).

Zur Versuchswiederholung wurden 63 (60)¹⁾ Vpn herangezogen, die nach Zufall (9 Schüler pro Klasse) aus der Gesamtstichprobe (Hauptversuche I; N=185) gezogen worden waren. Die Vpn wurden zu drei Versuchsgruppen zusammengefaßt, wodurch die Gruppengröße für die Verfahren ZA und ZS der durchschnittlichen Klassengröße (Hauptversuche I) annähernd entsprach. Die Zuordnung der drei V1²⁾ zu den Versuchsgruppen und den AN-Verfahren wurde wiederum systematisch variiert (zur Begründung siehe S. 62):

AN-Verfahren \ Versuchsgruppen	1	2	3
ZA	I ^{+))}	II	III
ZS	III	I	II
RW	II	III	I

+) V1 I bis III

Die Versuchswiederholung wurde zu Beginn des 6. Schuljahres im Oktober 1976 (2. Monatshälfte) vorgenommen. Die Bedingungen der ersten und der wiederholten Versuchsdurchführung entsprachen einander. Lediglich zwei Änderungen wurden vorgenommen. In die Instruktionen wurde - zur Begründung der Versuchswiederholung gegenüber den Vpn - ein entsprechender Vorspann aufgenommen.

-
- 1) Von den ursprünglich gezogenen 63 Vpn fehlten 3 Vpn bei mindestens einem der 3 AN-Verfahren, so daß für 60 Schüler auswertbare Daten vorlagen.
 - 2) Die Versuche wurden von der Verfasserin zusammen mit einem Studenten und einer Studentin durchgeführt.

Außerdem mußten für die Gruppenverfahren die Zeitbegrenzungen verkürzt werden, um einen zu erwartenden Übungsgewinn aufzufangen.

Im Hinblick auf die Qualität der Ergebnisse erschien es uns wichtig, den Kindern eine - möglichst plausible - Begründung für die Versuchswiederholung zu geben. Beim Zahlenaddieren bzw. Zahl-Symbol-Substituieren¹⁾ war zu Beginn der Instruktion folgendes eingefügt:

"Wie Ihr sicher alle schon mal erlebt habt, hat man manchmal einen guten Tag und manchmal einen schlechten. Manchmal könnt Ihr mehr leisten und manchmal weniger. Das hängt davon ab, wie man sich gerade fühlt.

Genauso kann es Euch natürlich bei den Aufgaben ergangen sein, die Ihr im letzten Schuljahr vor den Ferien bei uns gemacht habt. Deshalb wollen wir jetzt diese Aufgaben noch einmal wiederholen."

Der weitere Wortlaut der Instruktion blieb unverändert. Bei dem sich am folgenden Tag anschließenden Verfahren wurde lediglich auf die bereits gegebene Begründung verwiesen, ebenso beim zuletzt durchgeführten Ringewerfen:

"Ihr habt ja nun gestern ("... (Name) ..., Du hast ja bei den anderen Aufgaben ...") schon die Gründe gehört, warum wir diese Aufgaben wiederholen wollen. Darum können wir jetzt gleich anfangen."

1) Die Verfahren ZA und ZS wurden wiederum in alternierender Folge durchgeführt.

Die pro Durchgang vorgegebene Zeit wurde beim ZA von 35 auf 28 Sekunden, beim ZS von 18 auf 15 Sekunden reduziert. Für die Bestimmung der neuen Zeitbegrenzungen wurden zum einen die Durchschnittsleistungen in den Hauptversuchen I (bezogen auf die jeweiligen Zeitgrenzen), zum anderen die in den Vorversuchen ermittelten durchschnittlichen Leistungszuwächse von der ersten zur zweiten Messung (Zeitabstand ebenfalls 3 1/2 Monate) berücksichtigt.

4. Darstellung und Diskussion der Ergebnisse

Zur statistischen Auswertung der Daten werden durchweg nichtparametrische Verfahren eingesetzt. Damit sollen die im Fall parametrischer Verfahren möglichen Einwände in bezug auf Skalierungs- und Verteilungsvoraussetzungen aufgefangen werden.

4.1. Beschreibung der Ausgangsdaten

Auf den folgenden Seiten sind die Ausgangsdaten der Untersuchung (Hauptversuche I; N = 185) in einer Gesamttabelle (Tabelle 6) zusammengestellt. Die Testwerte im FAT 4-6 werden als Rohwerte mitgeteilt. Auf eine Transformation der Rohwerte in Prozentränge bzw. T-Werte wird verzichtet, da sich eine Anwendung der Klassennormen erübrigt (die Stichprobe setzt sich nur aus Schülern einer Klassenstufe zusammen) und die außerdem vorhandenen schulspezifischen Normen - bei einer Gesamtschulstichprobe - ebenfalls nicht indiziert sind. Angesichts des umfangreichen Datenmaterials aus den AN-Versuchen werden die AN-Daten nicht als Rohwerte, sondern nur in verrechneter Form dargestellt. Die als Maß für das AN verwendete individuelle Zieldiskrepanz (ZD) ist definiert als die mittlere Differenz zwischen jeweils erreichter und unmittelbar anschließend vorgenommener Leistung:

$$ZD = \frac{\sum_{i=1}^n (AN_{i+1} - LN_i)}{n}$$

AN: vorgenommene Leistung

LN: erreichte Leistung

n: Anzahl der Versuchsdurchgänge

Der ZD-Wert ist positiv, wenn das mittlere Zielniveau über dem mittleren Leistungsniveau liegt, Null, wenn beide einander entsprechen und negativ, wenn das mittlere Zielniveau unter dem mittleren Leistungsniveau liegt. - In motivationspsychologischer Sicht erscheint es ertragreicher, das AN nicht als ein absolut, sondern als ein relativ definiertes Ziel zu verstehen, als eine auf das tatsächlich erreichte Leistungsniveau bezogene Zielabweichung (Zieldiskrepanz). - Es sind ferner die jeweils über alle Durchgänge ($n = 11$) gemittelten Anspruchs- und Leistungsniveauwerte berechnet worden ($\overline{AN}$, $\overline{LN}$).¹⁾ Die Zeugnisnoten des 1. und 2. Halbjahres werden zunächst in unverrechneter Form dargestellt. (Die zur weiteren Verrechnung verwendeten Notendurchschnitte finden sich unter Punkt 4.1.3. in Tabelle 12.) Zusätzlich sind für die einzelnen Kennwerte die jeweiligen Gesamtmittelwerte und -standardabweichungen (für $N = 185$) in die Tabelle aufgenommen (mit Ausnahme der $\bar{X}$ - und s -Werte für die Noten; diese sind in Tabelle 11 angegeben).

1) Die zusätzlich erhobenen Urteile über erlebten Erfolg bzw. Mißerfolg (vgl. S. 46 ff.) und wiesen kaum eine Differenzierung auf und wurden in die im weiteren dargestellte Auswertung nicht einbezogen.

Tabelle 6:

Geschlecht (G: 1:männlich, 2:weiblich), Testwert im FAT 4-6 (RW: Rohwert). ZD-, AN- und LN-Werte jeweils für die 3 AN-Verfahren Zahlenaddition, Zahl-Symbol-Substitution, Ringewerfen. Zeugnisnoten im Zwischenzeugnis (1:Halbjahr) und im Jahreszeugnis (2:Halbjahr). Im 2.Halbjahr sind zusätzlich für die 3 Differenzierungsfächer (Deutsch, Englisch, Mathe.) die jeweiligen Kurseinstufungen (1: A-Kurs, 2: B-Kurs, 3: C-Kurs, 4: Schüler mit Sonderbetreuung) aufgenommen.

Vp	G	FAT RW	ZAHLENADDITION				ZAHL.-SYMB.-SUBST.				RINGEWERFEN				Z E U G N I S S E N			
			ZD	AN	LN	ZD	AN	LN	ZD	AN	LN	1. HALBJAHR	2. HALBJAHR	KURS				
												RDMEPBMKHS	RDMEPBMKHS	DEM				
1	2	50	-5	3	15,64	21,64	-4	2	9,64	14,64	-1	3	6,00	7,55	1111111222	2211111222	111	
2	2	57	-6	2	10,45	16,64	-7	7	5,55	13,27	-3	2	2,91	6,00	2222222222	2354222222	111	
3	2	44	-0	4	9,91	10,09	-0	6	9,91	11,45	+0	4	6,27	5,91	4455333222	3444333222	333	
4	1	45	+0	2	16,18	16,55	-2	8	11,45	15,27	-0	5	6,55	7,09	35564342323	42221232332	444	
5	1	40	+1	1	12,09	10,18	+0	4	10,00	9,82	+0	7	6,82	6,64	56645465344	52211245335	444	
6	2	66	+1	0	12,73	11,64	+1	9	13,36	11,91	+1	3	7,36	6,64	23343342234	33343332334	222	
7	2	55	-0	1	12,00	12,82	+0	2	14,09	14,73	+0	9	5,45	4,36	33342443333	34553433234	222	
8	2	57	-0	7	10,73	11,64	-0	8	12,82	14,00	+0	0	5,36	5,55	33343233233	44353323334	222	
9	2	65	-0	8	11,64	12,45	-2	9	15,45	19,36	+0	6	7,45	6,73	34343322231	24453321331	222	
10	1	50	+1	2	15,09	14,36	+1	9	12,91	11,55	+1	8	11,27	9,36	44334442321	23453443322	322	
11	2	45	-0	6	9,64	10,18	-0	5	11,73	12,73	+1	0	9,18	8,18	44545443231	35324433221	233	
12	1	49	+1	0	13,55	13,36	+1	3	12,18	11,73	+1	2	8,82	7,55	35645453233	33325343332	333	
13	2	66	-0	4	13,27	14,64	-0	3	11,36	12,18	+0	3	6,27	5,82	23221222331	23341222221	211	
14	2	74	+1	9	12,55	9,82	+0	4	10,18	10,27	+1	1	7,27	6,09	33455453235	34244443234	233	
15	1	17	+3	1	10,91	7,82	+1	1	10,09	8,82	+1	7	7,18	5,55	36666444423	42343345433	444	
16	1	45	+1	3	12,09	9,73	+2	0	9,45	7,45	+1	7	9,27	7,82	44555463233	55425553232	233	
17	2	85	-3	0	12,45	15,64	-3	5	12,73	16,73	-1	3	5,36	6,82	21112222221	22312222221	111	
18	1	50	+0	1	9,09	9,45	-0	2	9,64	10,55	+1	1	7,73	6,91	35533343331	45343333322	332	
19	1	50	-2	4	8,55	11,45	-0	1	10,73	11,09	-0	9	5,73	6,73	45454533233	54344453325	333	
20	1	67	+0	7	9,64	9,18	-1	2	13,18	15,55	+1	3	8,18	6,45	43432443332	44252333332	232	
21	1	55	-3	6	15,73	20,27	-5	7	10,91	17,27	-1	6	7,45	9,09	3433234321	34443334321	222	
22	2	69	-3	0	15,82	19,82	-4	6	13,36	18,82	-2	4	5,00	7,55	12222112332	23322112333	211	

Z E U G N I S N O T E N																	
1. HALBJAHR 2. HALBJAHR																	
VP	G	FAT	ZD	ZAHLENADDITION				ZAHL.-SYMB. -SUBST.				RINGEWERFEN				KURS	
				AN	LN	ZD	AN	LN	ZD	AN	LN	RDMEPMKHS	1) RDMEPMKHS	DEM 2)			
23	2	82	+0.8	15,73	15,55	-0.3	14,55	15,73	6,73	6,18	332322433331	23442233322	212				
24	2	29	-0.6	13,64	14,45	-0.9	11,45	20,82	4,73	6,09	33233333341	23452333332	212				
25	1	44	-0.3	20,27	20,64	-0.9	17,73	19,27	8,18	9,00	34122233211	33342334322	211				
26	1	19	+1.8	11,82	10,00	-2.3	12,09	14,55	8,91	5,45	45666453423	42333344433	444				
27	2	50	+0.0	15,00	15,00	-5.7	5,91	11,73	7,45	7,64	25554343324	14234353423	333				
28	1	43	-1.6	8,82	10,91	-2.2	10,27	13,36	1,1	5,45	36664333433	22241233433	444				
29	1	27	+2.6	12,82	10,64	-0.1	8,36	8,91	1,2	8,82	45666542422	45555552422	333				
30	2	74	+2.2	19,73	18,18	+3.4	19,91	17,27	1,2	6,18	12221212224	13433222321	111				
31	1	81	+0.7	19,82	20,09	+2.7	17,18	15,09	1,0	8,73	12212112224	1232222324	111				
32	2	55	+0.4	16,36	16,45	-1.4	14,18	16,45	1,3	5,55	14333433323	13343433323	222				
33	2	23	+0.6	9,91	9,18	-0.7	11,18	12,18	1,0	6,45	35554352332	24354452332	333				
34	2	57	+0.3	13,64	13,55	-2.3	12,73	16,00	0,9	8,73	34354523323	3444442333	223				
35	2	58	-1.1	12,73	13,64	-0.1	10,00	10,18	1,3	7,45	23434323313	23554332312	222				
36	1	39	+0.1	16,27	16,09	-1.8	9,82	11,91	0,5	10,55	24433243334	24544343334	322				
37	2	39	-1.1	13,00	14,09	-2.7	10,36	13,73	0,6	8,27	24655542322	25335432322	233				
38	1	17	-3.8	10,82	15,45	-3.3	8,55	12,64	1,1	7,36	25554433322	25434432323	333				
39	2	74	-0.4	20,00	20,82	-0.8	15,91	17,18	0,1	8,64	13321311211	12232221211	221				
40	2	70	+0.1	28,91	28,82	+1.3	22,91	22,09	1,3	11,36	11111311211	11111211211	111				
41	2	73	+0.4	22,27	22,45	+2.4	18,82	17,09	1,9	7,36	12312221211	1222211222	121				
42	1	42	+1.5	20,82	20,18	+1.8	18,00	17,36	1,3	11,73	10,45	2452321131	14353331331	231			
43	2	35	-1.2	11,64	13,64	-0.2	11,55	12,27	0,2	8,27	24554442213	24234452213	233				
44	1	54	-1.7	10,91	13,36	-4.2	9,45	13,91	-1,1	7,91	9,27	45543322432	34324331421	333			
45	1	54	-4.3	10,64	15,45	-4.9	7,18	12,64	-1,0	6,18	7,45	45654453424	35424453425	333			
46	2	46	-1.5	10,09	12,00	-2.4	9,09	12,18	0,0	8,73	8,64	24444442331	23244442331	333			
47	2	57	-2.3	12,45	15,27	-0.4	12,00	13,09	1,5	10,45	9,73	23453332231	1442334223	223			
48	2	35	+0.8	12,82	12,18	+0.5	15,09	15,36	0,2	8,45	25543452334	25354352334	221				
49	1	60	-1.5	16,82	18,64	-1.8	11,36	13,64	0,4	11,55	11,36	13211212231	1223222222	221			
50	1	67	+2.3	17,36	15,45	+0.2	13,27	14,36	2,6	12,91	10,82	13433222412	13243331322	222			
51	1	17	-0.2	4,82	4,82	-3.1	7,18	11,09	12,91	10,82	46665564334	42622354435	444				
52	1	59	-0.3	13,82	14,00	-1.5	8,45	9,91	0,4	8,18	7,73	23222223343	24323333343	212			

Vp	G	FAT RW	ZAHLENADDITION			ZAHL.-SYMB.-SUBST.			RINGEWERFEN			Z E U G N I S N O T E N			KURS DEM 2)		
			ZD	AN	LN	ZD	AN	LN	ZD	AN	LN	1. HALBJAHR 2. HALBJAHR					
												RDEMEPBKHS	RDEMEPBKHS				
53	1	52	+2.3	18,55	16,82	-2.3	12,09	14,55	+3.5	12,27	9,09	44,44	33,44	32,3	33,52	33,44	22,3
54	1	52	-0.4	15,55	15,91	-3.9	9,91	14,36	+0.2	6,45	6,09	14,32	44,21	11	13,54	23,34	22,1
55	1	77	-0.1	18,00	18,73	+1.1	12,36	12,36	+0.9	7,91	6,91	22,13	42,33	30	12,23	22,32	11,2
56	2	54	-1.6	15,09	17,73	-1.9	12,45	14,55	+0.4	4,82	4,45	23,35	43,42	30	23,42	33,34	22,3
57	1	79	+0.7	17,36	17,55	-0.1	17,27	17,82	+3.0	8,18	5,00	13,22	33,32	22	13,32	22,22	21,1
58	2	23	-1.7	10,82	12,73	-2.1	10,00	12,09	-0.9	6,73	8,18	55,55	45,43	40	44,34	35,44	33,3
59	2	25	+0.8	15,00	14,36	+0.9	12,36	11,00	+2.0	8,64	6,73	44,45	45,33	30	34,44	45,44	13,3
60	2	73	-2.3	17,64	20,73	-2.3	11,27	13,82	+0.8	7,64	7,00	12,11	21,22	20	13,32	12,23	11,1
61	1	66	+0.2	11,45	11,36	+1.7	10,55	8,73	+1.3	9,45	8,09	35,44	44,34	31	34,54	44,34	32,3
62	2	29	+0.5	10,91	11,36	+1.6	14,09	12,55	+1.1	9,09	8,27	55,46	35,43	30	44,46	55,44	33,3
63	2	57	-5.6	8,64	14,55	-0.6	10,18	11,09	-0.7	4,55	5,00	44,44	45,34	30	33,54	44,44	32,3
64	1	61	-3.4	9,45	13,36	-1.3	9,09	10,55	-0.7	5,00	6,09	13,34	33,42	23	14,53	33,34	22,2
65	2	58	-4.7	8,64	13,91	-2.2	12,27	15,18	-0.8	5,82	6,82	13,23	33,32	20	13,42	33,32	22,2
66	2	54	-12.2	9,00	22,09	-4.9	13,36	18,45	-1.3	6,27	7,64	13,33	34,22	20	12,45	23,22	22,2
67	2	37	-3.8	10,09	14,27	+0.9	16,91	15,91	+0.6	10,73	10,00	35,55	44,22	30	24,44	43,22	33,3
68	2	31	+0.4	9,64	9,18	+0.9	10,45	9,27	-0.2	5,09	5,55	35,55	44,33	30	34,65	55,44	23,3
69	1	54	+0.2	11,18	11,00	-0.3	12,82	13,18	+0.1	5,73	5,55	25,55	33,44	33	23,42	34,44	33,3
70	2	50	-2.8	6,73	9,91	-6.0	6,73	12,82	-1.4	5,73	7,27	23,46	54,33	30	24,65	44,34	23,3
71	2	49	-0.6	10,18	10,09	-1.4	11,00	12,73	+1.1	7,82	7,09	33,45	44,32	20	23,54	44,34	23,3
72	1	46	+0.7	16,00	16,09	+0.4	10,91	10,27	+0.9	11,73	10,91	25,44	44,32	1	25,36	44,34	13,2
73	2	16	-0.7	5,64	6,55	+1.5	13,64	12,91	+0.5	7,82	7,45	46,66	56,42	30	33,33	44,32	14,4
74	1	34	+1.2	12,55	12,36	+2.6	11,91	10,82	+2.8	10,73	7,82	25,55	33,34	33	25,54	33,44	33,3
75	2	25	-4.6	9,09	14,36	-3.4	9,27	13,18	-0.3	6,55	7,00	34,45	44,34	30	35,25	45,34	33,3
76	2	58	+0.2	16,00	16,18	-0.8	10,73	12,09	+0.2	6,82	6,82	44,44	44,33	30	33,34	44,44	33,3
77	2	21	+2.5	7,64	5,09	+1.3	12,27	11,36	+2.7	10,82	7,91	35,65	66,42	30	32,24	43,44	44,4
78	1	83	+1.3	24,36	24,64	+0.5	16,27	16,45	+0.6	9,55	9,09	11,11	11,13	21	12,21	12,14	11,1
79	2	67	+1.5	16,73	16,18	+1.1	13,00	12,82	+0.3	8,64	9,09	12,22	32,32	30	12,44	33,22	11,1
80	1	15	+0.8	17,64	17,91	-0.1	8,18	8,36	-0.3	7,91	8,18	35,46	55,44	23	34,55	55,44	23,3

VP	Z E U G N I S N O T E N																													
	ZAHLENADDITION					ZAHL.-SYMB.-SUBST.					RINGWERFEN					1) RDEMEPBKHS					KURS									
	G	FAT	ZD	AN	LN	ZD	AN	LN	ZD	AN	LN	ZD	AN	LN	ZD	AN	LN	ZD	AN	LN	ZD	AN	LN	ZD	AN	LN	ZD	AN	LN	
81	1	51	+0 2	10,82	10,91	+0 6	8,82	8,73	+1 5	12,09	10,55	24334333313	23434433223	200																
82	2	68	-2 2	14,09	17,36	+0 4	11,09	11,64	+0 0	7,64	7,73	13322212320	12342212214	201																
83	2	76	-3 7	15,82	20,55	-2 4	10,09	13,18	-1 1	5,73	7,09	13112312220	12322322113	211																
84	2	43	-0 7	15,09	16,82	-2 6	8,00	11,00	+0 2	6,73	6,45	1333443420	13354343322	202																
85	2	57	-4 7	6,64	11,27	-4 5	6,36	11,27	+0 3	9,00	9,91	4555444330	44345444335	212																
86	1	82	-2 8	8,45	12,18	-3 1	6,36	9,82	+0 3	6,09	6,00	2322322222	23343222222	203																
87	2	36	-0 2	8,73	9,36	-0 1	9,09	9,73	-1 4	2,55	4,27	4455454340	4435554355	213																
88	1	68	-0 5	10,00	11,00	-1 5	8,27	10,55	+1 1	9,82	9,27	24332233313	12322222224	111																
89	1	80	+1 4	11,55	10,64	-1 7	8,09	10,18	-2 2	6,45	8,82	12221132323	12322222224	111																
90	2	49	-3 4	5,27	14,36	-6 8	5,27	12,82	-0 7	4,91	5,45	23354332440	23354332442	204																
91	1	85	-1 6	16,27	18,55	-2 8	7,82	11,00	-0 3	10,55	11,45	1233232222	12433332321	204																
92	1	47	-0 5	12,18	12,91	+0 2	9,64	10,00	+1 6	11,27	9,64	25654454332	24455444232	205																
93	1	11	-1 0	10,91	12,09	-2 3	6,73	9,36	+2 3	7,73	5,09	35654543432	34436554432	205																
94	1	19	+1 5	10,45	9,09	+1 1	8,82	8,36	+1 8	9,73	7,82	35665454333	35455454222	206																
95	1	78	+0 9	20,09	19,73	+1 5	12,91	12,55	+1 4	12,73	11,36	13322323221	12332332321	206																
96	1	62	+1 5	16,64	14,45	+1 6	11,00	10,00	+1 0	4,64	3,73	22432323223	22332223233	102																
97	2	62	+0 4	14,18	14,73	+0 1	12,00	12,18	+0 0	10,45	10,64	13224442320	1344442323	215																
98	2	67	+0 7	8,27	7,82	+1 2	9,09	8,17	-0 1	11,00	8,36	33444332420	23444432223	207																
99	2	55	+0 7	13,91	14,18	+1 5	12,82	12,18	+2 7	11,00	8,64	24533323430	14344333332	207																
100	2	51	+0 1	9,73	9,64	-0 5	7,64	8,64	+0 1	6,45	6,18	14434332420	13244333322	208																
101	2	51	-1 3	10,27	12,64	-2 2	8,45	11,18	-1 0	5,27	6,64	24454442340	23444443242	208																
102	2	48	+1 3	18,09	18,09	+2 2	13,27	11,91	-0 5	10,64	10,82	23333441430	23254423336	209																
103	2	43	+2 6	14,09	12,09	+1 4	11,91	10,36	+1 7	10,73	9,27	13333232330	13353333332	209																
104	2	50	-1 0	8,91	10,00	-1 4	7,82	8,64	+0 0	4,55	4,55	3445543440	34545533333	210																
105	1	49	+0 4	11,27	11,45	-0 5	12,27	13,09	+0 5	9,45	9,09	44554453332	34454453332	210																
106	1	67	+2 2	12,55	10,73	+0 8	14,45	14,00	+0 3	11,18	10,64	34522323232	24334322332	209																
107	2	61	-0 2	12,82	12,91	-5 0	10,00	15,09	-1 7	7,45	9,45	23454353430	23433453333	210																
108	1	75	-0 7	5,82	6,09	-1 5	7,27	8,55	-0 3	6,36	6,82	23354454333	23254444323	211																

Vp	Z E U G N I S N O T E N										KURS		
	1. HALBJAHR 2. HALBJAHR										DEM 2)		
	Z A H L E N A D D I T I O N										R D E M E P B M K H S 1)		
G	F A T		Z A H L - S Y M B . - S U B S T .		R I N G E W E R F E N		R I N G E W E R F E N		R D E M E P B M K H S 1)		R D E M E P B M K H S		
	RW	RD	AN	LN	RD	AN	LN	RD	AN	LN	RD	AN	
109	1 78	-1 5	10,45	12,27	-1 1	9,18	10,55	+0 6	8,73	8,00	121233123321	11232112321	111
110	1 74	-1 8	10,73	12,82	+0 4	12,73	12,73	+0 5	11,64	11,45	14324322323	23323222224	221
111	2 70	-5 2	9,18	14,36	-7 6	8,27	16,18	-1 9	7,45	9,45	12223223330	12332123231	112
112	1 43	+1 1	16,18	15,18	+0 0	12,64	13,18	+0 3	10,91	10,91	14534323224	23323222225	332
113	2 54	-0 4	10,91	11,82	+0 0	9,27	9,82	-0 3	7,55	8,00	24144233330	23344332334	212
114	1 63	+0 8	11,45	9,82	+1 4	11,91	10,55	+0 4	8,82	8,55	24544324421	23344322321	332
115	2 52	-0 5	10,64	10,82	+0 6	11,82	11,55	+1 4	9,82	8,45	23354434330	22425432225	223
116	1 67	+0 9	15,55	14,45	+1 2	11,82	11,09	+0 9	6,45	5,36	23334223431	23445333332	222
117	1 73	-2 0	20,18	23,00	+0 8	14,73	14,27	-0 1	10,00	10,09	13212122324	12322222324	211
118	1 55	-0 7	7,09	7,82	-0 5	8,82	10,00	+0 9	4,00	3,36	35564443322	34355442322	333
119	2 69	-2 3	9,55	11,45	+0 3	11,55	11,45	+0 4	5,55	5,27	13343332320	23334331325	222
120	2 69	-1 2	16,09	18,00	-0 9	12,73	13,91	-0 2	5,27	5,45	12232223230	12323221223	111
121	1 33	-0 2	6,45	6,82	+1 5	9,64	8,27	+0 9	7,82	6,82	35665454442	35455454333	333
122	2 70	+0 6	15,27	15,18	-1 1	11,09	12,64	+0 1	6,00	5,91	13224232320	12333222216	211
123	2 74	+0 9	16,91	16,82	+2 1	14,00	12,82	+2 3	12,09	10,18	12134233420	122342231322	112
124	2 70	-1 3	10,64	11,82	-0 8	9,91	11,09	+0 0	8,82	9,18	23223222420	21223222326	211
125	1 42	+0 2	12,27	11,36	-0 3	9,09	9,55	+1 2	10,36	9,18	13334323433	23434322334	222
126	1 55	+0 8	9,64	9,09	-1 5	10,18	11,91	+3 2	12,82	9,27	15666434320	24456433222	333
127	2 80	-3 6	11,45	15,45	-3 0	9,27	12,82	+0 1	7,91	7,36	11112121210	11212121213	111
128	1 76	-0 6	13,27	13,82	+0 1	9,91	10,27	+0 4	8,36	7,82	23345344331	22434343222	222
129	1 49	+1 7	16,55	15,91	+2 2	14,73	13,18	+1 3	6,91	5,55	34564444430	23245543333	333
130	2 41	+0 4	10,55	10,09	+0 7	14,55	14,55	+1 4	11,36	10,18	24565343220	24455443222	333
131	1 54	-1 4	11,82	13,09	+0 2	10,73	10,27	-0 2	9,55	9,64	24435354333	24444343332	222
132	1 41	+1 9	10,64	8,45	-0 3	9,64	10,18	+0 0	6,18	5,55	24445444442	24435333330	223
133	2 55	-0 6	11,55	12,55	+0 2	11,73	11,73	+0 1	4,55	4,55	15665343320	24455453321	333
134	1 71	+0 1	13,36	13,91	-0 6	8,73	9,73	+0 2	5,82	5,73	13323343321	14423232222	221
135	2 36	+0 3	9,27	9,00	+1 0	9,55	8,64	+0 4	9,27	9,00	23444434230	23454544222	223
136	2 75	-3 5	10,82	14,55	-2 8	10,00	13,82	-1 8	3,64	5,64	13212233320	2342332214	211

VP	Z E U G N I S N O T E N										KURS	
	1. HALBJAHR					2. HALBJAHR					DEM 2)	
	G	FAT	ZD	AN	LN	ZD	AN	LN	AN	LN	RDMEPBKHS ¹⁾	RDMEPBKHS
		RW										
137	2	23	+2.4	10,18	8,27	+0.4	9,82	10,09	8,27	7,55	355545553350	35354654346
138	1	38	+0.2	12,00	12,55	+0.4	11,27	11,64	9,00	8,91	344445552225	33454555225
139	1	70	-1.1	13,73	14,64	-2.1	12,45	15,64	9,91	9,45	23332242224	23333242214
140	2	52	-5.1	10,45	16,18	-5.8	9,09	15,82	5,27	4,91	13223112230	13543222211
141	1	40	-0.7	12,00	13,36	+1.7	11,45	10,55	6,91	5,91	35645543334	35465643333
142	1	67	-3.1	17,91	22,00	-4.0	11,36	16,09	10,00	10,36	14323444321	24453443320
143	1	35	+0.3	12,00	11,45	-0.6	10,73	10,36	7,55	6,64	35655543335	25455543324
144	2	80	-1.3	11,73	13,00	-1.3	10,09	11,73	5,18	6,27	13222234230	23442233225
145	2	74	-3.3	12,45	16,36	-3.1	10,55	14,27	7,18	7,91	12222323210	12342333214
146	2	66	+0.0	16,18	17,09	+0.9	13,73	13,82	8,82	7,91	12112322230	12242333222
147	2	54	+0.2	12,64	12,82	+0.4	13,09	13,55	4,64	4,36	23344424230	23333435222
148	2	50	+0.9	12,82	12,45	+0.7	12,82	12,73	8,55	7,91	34544543330	24353533323
149	1	77	-0.3	13,09	13,91	-0.9	8,27	9,55	10,09	11,82	12121222322	12221212224
150	2	46	-0.2	8,27	8,27	-1.8	9,45	11,91	7,45	8,00	45655563230	45465553221
151	2	30	+0.5	9,00	8,55	+0.5	8,73	9,00	8,73	8,82	45565554340	45365555335
152	2	72	-3.6	12,73	16,55	-1.8	10,55	13,00	8,73	9,91	12223322220	12353322311
153	2	33	-4.4	12,09	17,00	-2.8	11,18	14,82	6,18	7,55	35656564230	35455555236
154	1	42	-0.6	9,18	10,18	-2.0	10,00	12,55	8,45	7,64	34544434435	34454444434
155	2	66	-3.7	17,00	21,64	-2.5	11,27	14,36	9,55	9,73	11212113230	11232212224
156	1	69	-4.4	11,55	16,27	-7.4	3,45	10,91	2,09	6,91	23233333341	33653444331
157	1	53	-0.2	11,36	11,91	-6.5	3,82	10,27	3,91	8,09	25544343322	24354454333
158	2	44	-6.3	12,91	19,91	-3.3	10,09	14,45	7,91	9,36	13244553240	23444544235
159	2	60	-3.1	17,27	20,82	-5.5	8,64	15,09	7,55	8,91	14313553230	23342442225
160	1	55	+2.4	16,36	14,82	+1.3	13,00	12,82	11,27	10,36	24534453213	34244453222
161	1	57	-3.9	4,73	8,55	-3.5	4,64	8,36	6,73	7,55	34344364433	34544343433
162	2	89	-5.0	12,73	17,64	-2.3	10,36	13,18	3,18	3,82	11111122220	11111112214
163	2	71	-3.9	10,00	13,91	-1.7	10,27	12,09	9,55	10,73	12222221220	12222221222
164	1	68	-1.2	15,27	16,27	+0.1	12,91	13,36	8,18	8,73	24333243321	24433343222

4.1.1. Zur Variablen Intelligenz

Die im FAT 4-6 erzielten Testwerte (Rohwerte) sind in Tabelle 6 mitgeteilt. Die Häufigkeitsverteilung der FAT-Werte ist tabellarisch dargestellt (Tabelle 7). Die Gruppierung der Meßwerte erfolgt - hier und auch bei weiteren Häufigkeitsverteilungen - anhand folgender Kriterien: Die Anzahl der Klassen wird nach einer von STURGES (vgl. SACHS 1974, S. 46) empfohlenen Formel bestimmt. Die Größe bzw. Breite des Klassenintervalls sowie der Beginn des ersten und das Ende des letzten Intervalls werden nach einer von LEWIS vorgeschlagenen Formel (vgl. LIENERT 1973, S. 41) festgelegt. - Eine theoretisch einheitliche Lösung für Gruppierungsprobleme gibt es nicht. Die hier anfallenden Entscheidungen werden üblicherweise unter inhaltlichen Aspekten vom V1 getroffen, wobei diese Entscheidungen von bestimmten Erwartungen des Vls beeinflusst sein können. Um diesen möglichen Einflüssen vorzubeugen, wird über die bereits genannten Formeln entschieden.

Tabelle 7: Häufigkeitsverteilung der Testwerte (Rohwerte) im FAT 4-6 (N = 185)

Klassen- grenzen	Häufigkeiten	
	absolut	relativ (in %)
10 - 18	6	3.2
19 - 27	10	5.4
28 - 36	13	7.0
37 - 45	20	10.8
46 - 54	36	19.5
55 - 63	33	17.8
64 - 72	34	18.4
73 - 81	24	13.0
82 - 90	9	4.9

Dieser Häufigkeitsverteilung wird zum Vergleich die für die Eichstichprobe des FAT erhaltene Rohwertverteilung (dargestellt im Beiheft zum FAT 4-6; BELSER und RAATZ 1973, S. 12) gegenübergestellt. Dazu wird die Gruppierung unserer Daten an die von BELSER und RAATZ gewählte Klasseneinteilung angepaßt.

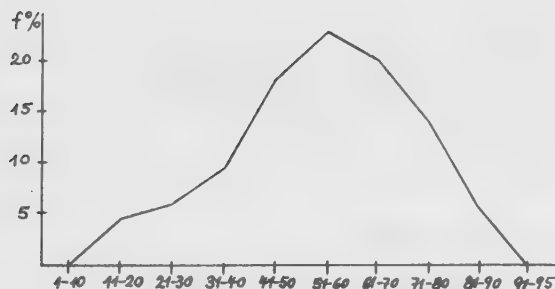


Abbildung 1: Prozentuale Verteilung der
FAT-Rohwerte (N = 185)

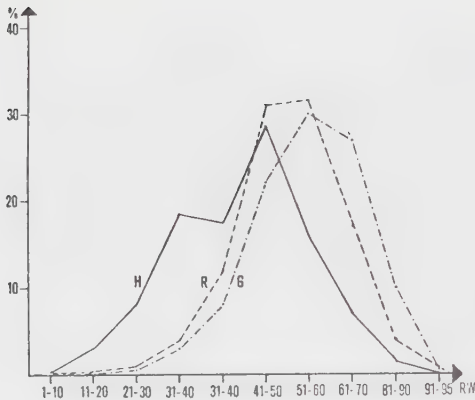


Abbildung 2: Prozentuale Verteilung der FAT-Rohwerte für die 5. Klassen der Eichstichprobe (N = 6002; 1986 Hauptschüler (H), 1657 Realschüler (R), 2359 Gymnasiasten (G)) (nach BELSER und RAATZ 1973, S. 12)¹⁾

4.1.2. Zur Variablen Anspruchsniveau

Die Häufigkeitsverteilungen der ZD-Werte für die drei AN-Verfahren (ZA, ZS und RW) sind in den Tabellen 8, 9 und 10 dargestellt und in den Abbildungen 3, 4 und 5 veranschaulicht. Die jeweils zugehörigen ZD-Werte sind der Tabelle 6 (auf den Seiten 81 bis 87) zu entnehmen.

1) Die Klasse 31-40 ist im Original doppelt aufgeführt, dafür fehlt die Klasse 71-80.

Tabelle 8: Häufigkeitsverteilung der ZD-Werte für
ZA (N = 185)

Klassen- grenzen	Häufigkeiten	
	absolut	relativ (in %)
-12.60 - -10.83	1	0.5
-10.82 - - 9.05	1	0.5
- 9.04 - - 7.27	0	0.0
- 7.26 - - 5.49	4	2.2
- 5.48 - - 3.71	19	10.3
- 3.70 - - 1.93	23	12.4
- 1.92 - - 0.15	56	30.3
- 0.14 - 1.63	65	35.1
1.64 - 3.41	16	8.6

Tabelle 9: Häufigkeitsverteilung der ZD-Werte für
ZS (N = 185)

Klassen- grenzen	Häufigkeiten	
	absolut	relativ (in %)
-9.35 - -7.91	1	0.5
-7.90 - -6.46	5	2.7
-6.45 - -5.01	6	3.2
-5.00 - -3.56	12	6.5
-3.55 - -2.11	32	17.3
-2.10 - -0.66	33	17.8
-0.65 - 0.79	54	29.2
0.80 - 2.24	37	20.0
2.25 - 3.69	5	2.7

Tabelle 10: Häufigkeitsverteilung der ZD-Werte für
RW (N = 185)

Klassen- grenzen	Häufigkeiten	
	absolut	relativ (in %)
-4.45 - -3.57	2	1.1
-3.56 - -2.68	1	0.5
-2.67 - -1.79	6	3.2
-1.78 - -0.90	24	13.0
-0.89 - -0.01	26	14.1
0.00 - 0.88	56	30.3
0.89 - 1.77	54	29.2
1.78 - 2.66	9	4.9
2.67 - 3.55	7	3.8

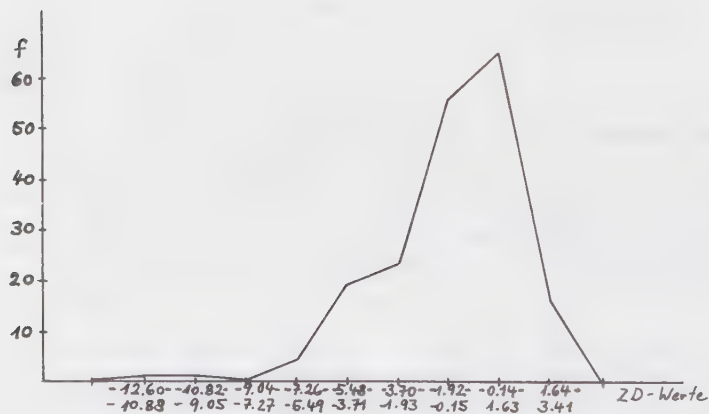


Abbildung 3: Häufigkeitsverteilung der ZD-Werte für
ZA (N = 185)

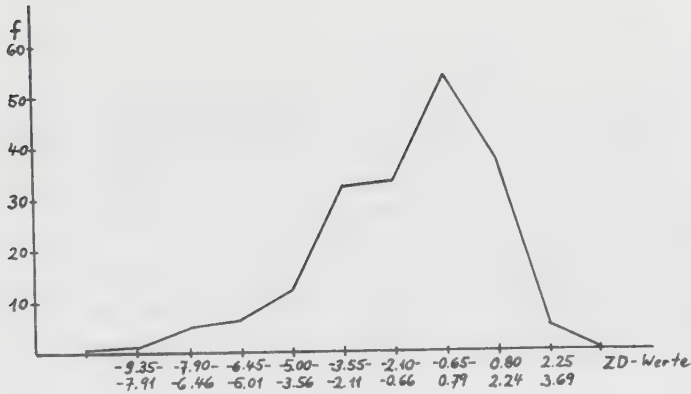


Abbildung 4: Häufigkeitsverteilung der ZD-Werte für ZS
(N = 185)

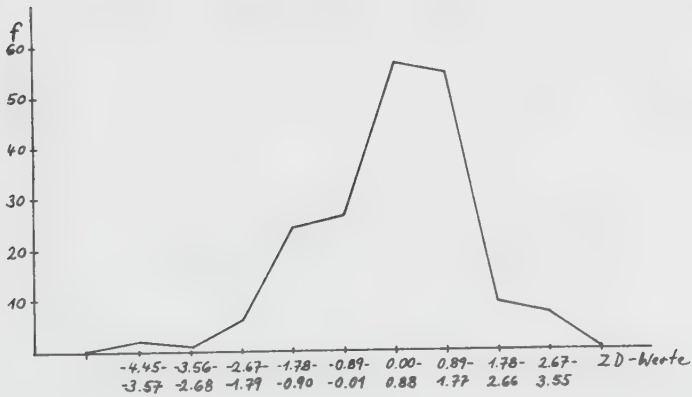


Abbildung 5: Häufigkeitsverteilung der ZD-Werte für RW
(N = 185)

Die Abbildungen 6, 7 und 8 stellen die durchschnittlich vorgenommenen ($\overline{AN}$) und die durchschnittlich erreichten ($\overline{LN}$) Leistungen in ihrem Verlauf über die Versuchsdurchgänge dar. Dabei sind - analog dem ZD-Maß - die $\overline{LN}$ -Werte den $\overline{AN}$ -Werten des jeweils nachfolgenden Durchgangs zugeordnet und auf gleichen Punkten der Abszisse aufgetragen. (Die Mittelwerte und Standardabweichungen der AN- und LN-Werte für die Verfahren ZA, ZS und RW sind der Tabelle 1 des Anhangs zu entnehmen.)

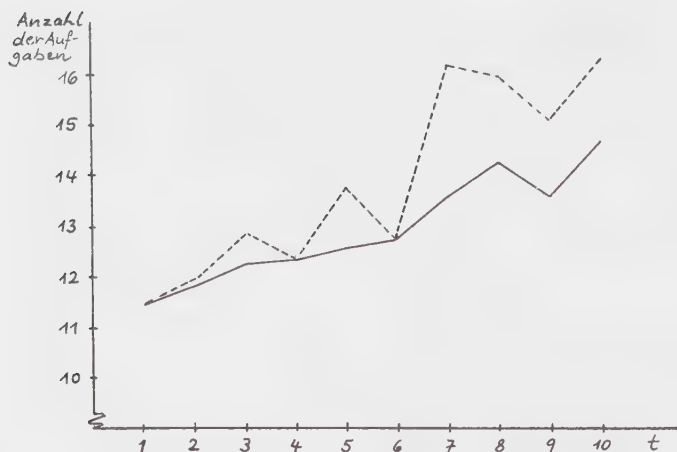


Abbildung 6: Verlauf der mittleren AN-Werte (—) und LN-Werte (---) für ZA (N = 185)

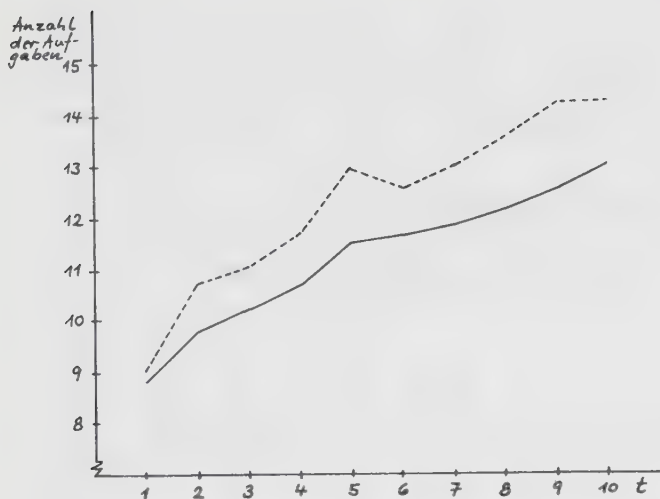


Abbildung 7: Verlauf der mittleren AN-Werte (—) und LN-Werte (---) für ZS (N = 185)

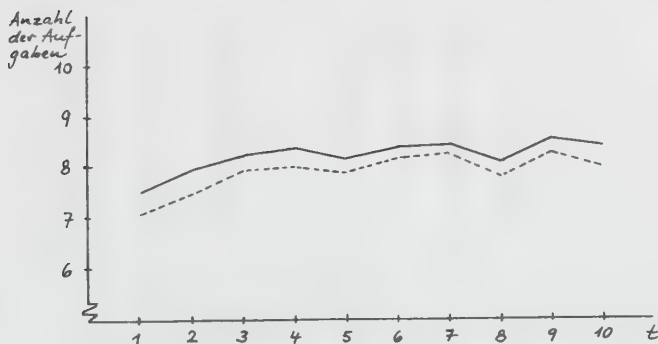


Abbildung 8: Verlauf der mittleren AN-Werte (—) und LN-Werte (---) für RW (N = 185)

4.1.3. Zur Variablen Schulleistung

Die Zeugnisnoten des 1. und 2. Halbjahres sind in Tabelle 6 mitgeteilt. Tabelle 11 zeigt die Häufigkeiten der einzelnen Notenstufen sowie die Mittelwerte und Standardabweichungen pro Fach jeweils für das 1. und 2. Halbjahr.

Tabelle 11: Häufigkeiten der Noten und $\bar{X}$ - und s-Werte pro Fach, getrennt für das 1. und 2. Halbjahr

	Fach	Notenstufen						Σ	$\bar{X}$	s
		1	2	3	4	5	6			
1. Halbjahr	Religion	61	59	41	21	3	0	185	2,17	1,06
	Deutsch	9	22	57	52	39	6	185	3,58	1,18
	Englisch	18	32	43	34	37	21	185	3,56	1,51
	Mathematik	18	37	34	41	36	19	185	3,52	1,50
	Erdkunde	10	35	45	64	25	6	185	3,42	1,19
	Physik/Chemie	14	38	61	49	21	2	185	3,17	1,14
	Biologie	12	42	43	55	27	6	185	3,33	1,25
	Musik	16	57	78	32	2	0	185	2,71	0,89
	Kunst	2	68	82	33	0	0	185	2,79	0,74
	Werken/Handarb.	15	78	78	13	1	0	185	2,50	0,77
	Sport	32	31	31	12	4	0	110 ^{*)}	2,32	1,12
2. Halbjahr	Religion	54	78	35	15	3	0	185	2,11	0,97
	Deutsch	6	41	58	60	19	1	185	3,26	1,04
	Englisch	4	34	73	58	13	3	185	3,28	0,97
	Mathematik	8	32	43	54	41	7	185	3,59	1,24
	Erdkunde	10	42	47	57	28	1	185	3,29	1,15
	Physik/Chemie	9	47	60	47	20	2	185	3,15	1,10
	Biologie	10	39	59	51	26	0	185	3,24	1,10
	Musik	18	69	60	32	6	0	185	2,67	0,98
	Kunst	1	73	82	29	0	0	185	2,75	0,72
	Werken/Handarb.	16	102	62	4	1	0	185	2,31	0,68
	Sport	35	52	40	32	15	9	183 ⁺⁺⁾	2,82	1,40

+) Wegen Unterrichtsausfall wurde für die Mädchen aus 5 Klassen keine Sportnote vergeben. Außerdem fehlten die Sportnoten von 2 Jungen, die vom Sportunterricht befreit worden waren.

++) Wegen Befreiung vom Unterricht lagen für 2 Jungen keine Sportnoten vor.

Die Noten in den einzelnen Fächern werden für die weitere Datenverrechnung (siehe unter Punkt 4.4.) zu bestimmten Notendurchschnitten zusammengefaßt. Pro Vp werden insgesamt sieben verschiedene Notendurchschnitte berechnet (siehe unten Tabelle 12). Mit dieser Form der Zusammenfassung schien uns eine genügende Differenzierung der Schulleistung gegeben zu sein, zumal es Hinweise gibt, daß die - über Noten erfaßte - Schulleistung sich keineswegs als so differenziert erweist, wie dies Zeugnisse mit entsprechend vielen (in der Regel etwa zehn) Einzelnoten nahelegen.¹⁾

1) Dies zeigte sich in faktorenanalytischen Untersuchungen zur Schulleistung (z. B. LANGFELDT und FINGERHUT 1974).

Tabelle 12: Notendurchschnitte pro Vp (N = 185)
(Die für die Notendurchschnitte verwendeten Abkürzungen sind am Ende der Tabelle erläutert)

Vp	NK 1	NK 2	NK12	GN 1	GN 2	KN 1	KN 2
1	1.37	1.50	1.43	1.27	1.45	1.00	1.33
2	2.00	2.00	2.00	2.11	2.67	2.00	4.00
3	2.75	2.62	2.68	3.46	3.24	4.66	4.00
4	3.00	2.50	2.75	3.95	2.65	5.33	2.00
5	4.50	3.50	4.00	5.08	3.24	5.33	1.66
6	2.87	3.00	2.93	3.46	3.38	3.33	3.33
7	3.12	3.12	3.12	3.49	3.85	3.33	4.66
8	2.75	3.12	2.93	3.22	3.71	3.33	4.00
9	2.37	2.25	2.31	3.02	3.15	3.66	4.33
10	3.00	2.87	2.93	3.36	3.46	3.33	4.00
11	3.25	2.75	3.00	3.85	3.22	4.33	3.33
12	3.50	3.25	3.37	4.25	3.38	5.00	2.66
13	2.00	1.75	1.87	2.47	2.48	2.33	3.33
14	3.75	3.37	3.56	4.04	3.58	4.00	3.33
15	3.75	3.62	3.68	4.73	3.78	6.00	3.00
16	3.75	3.75	3.75	4.43	4.07	4.66	3.66
17	1.87	1.87	1.87	2.03	2.27	1.00	2.00
18	2.87	2.87	2.87	3.45	3.38	4.33	4.00
19	3.50	3.87	3.68	4.13	4.03	4.66	3.33
20	3.12	2.87	3.00	3.55	3.45	3.33	3.66
21	2.62	2.75	2.68	3.14	3.40	3.33	4.00
22	1.87	2.12	2.00	2.19	2.58	2.00	2.66
23	2.62	2.37	2.50	2.83	2.96	2.66	3.66
24	2.87	2.62	2.75	3.07	3.26	2.66	4.00
25	2.12	2.75	2.43	2.46	3.20	2.33	3.33
26	3.87	3.50	3.68	4.58	3.56	5.66	2.66
27	3.12	3.12	3.12	4.05	3.41	5.00	3.00
28	3.25	2.62	2.93	4.36	2.94	6.00	2.66
29	3.62	3.62	3.62	4.57	4.26	5.66	5.00
30	2.00	2.00	2.00	2.41	2.75	2.00	3.33
31	1.87	2.25	2.06	2.03	2.52	1.66	2.33
32	2.75	2.62	2.68	3.09	3.04	3.33	3.33
33	3.12	3.12	3.12	3.91	3.64	5.00	4.00
34	3.12	3.25	3.18	3.71	3.78	4.00	4.00
35	2.62	2.50	2.56	3.15	3.34	3.33	4.33
36	3.00	3.25	3.12	3.46	3.84	3.66	4.33
37	3.12	2.87	3.00	3.95	3.44	5.00	3.66
38	2.87	2.87	2.87	3.81	3.49	5.00	4.00
39	1.37	1.50	1.43	2.07	2.04	2.66	2.33
40	1.37	1.25	1.31	1.46	1.36	1.00	1.00
41	1.50	1.62	1.56	1.76	1.85	2.00	2.00
42	1.87	2.25	2.06	2.52	2.89	3.66	4.00
43	2.75	2.87	2.81	3.50	3.17	4.66	3.00
44	2.87	2.62	2.75	3.68	3.01	4.66	3.00
45	3.75	3.75	3.75	4.51	4.00	5.33	3.66
46	2.75	2.87	2.81	3.50	3.27	4.00	3.00
47	2.37	2.50	2.43	3.13	3.02	4.00	3.33
48	3.25	3.25	3.25	3.92	3.82	4.66	4.33
49	1.62	1.87	1.75	2.08	2.34	2.00	2.33
50	2.12	2.25	2.18	2.64	2.66	3.33	3.00

Fortsetzung Tabelle 12:

Vp	NK 1	NK 2	NK12	GN 1	GN 2	KN 1	KN 2
51	4.25	3.75	4.00	4.96	3.87	6.00	3.33
52	2.62	2.87	2.75	2.99	3.35	2.33	3.33
53	3.25	3.25	3.25	3.72	3.57	4.00	3.33
54	2.37	2.37	2.37	2.88	3.14	3.00	4.00
55	2.75	2.12	2.43	2.80	2.46	2.00	2.33
56	2.87	2.62	2.75	3.34	2.95	3.66	3.00
57	2.25	1.87	2.06	2.57	2.45	2.33	3.00
58	3.87	3.75	3.81	4.41	3.95	5.00	3.66
59	3.87	3.87	3.87	4.40	4.26	4.33	4.00
60	2.00	2.12	2.06	2.21	2.75	1.33	3.00
61	3.37	3.25	3.31	3.83	3.79	4.33	4.33
62	3.37	3.62	3.50	4.16	4.25	5.00	4.66
63	3.87	3.75	3.81	4.28	4.11	4.00	3.66
64	2.87	2.75	2.81	3.38	3.55	3.33	4.33
65	2.12	2.12	2.12	2.58	2.86	2.66	3.66
66	2.12	2.00	2.06	2.59	2.71	3.00	3.66
67	2.87	2.75	2.81	3.69	3.33	5.00	4.00
68	3.50	3.75	3.62	4.24	4.39	5.00	5.00
69	3.25	3.12	3.18	4.11	3.67	5.00	3.66
70	3.62	3.75	3.68	4.19	4.33	4.33	4.66
71	3.12	3.12	3.12	3.74	3.57	4.00	3.33
72	3.00	3.00	3.00	3.70	3.77	4.33	4.66
73	3.75	2.87	3.31	4.70	3.25	6.00	3.00
74	3.00	3.25	3.12	3.97	3.93	5.00	4.66
75	3.37	3.50	3.43	3.99	3.99	4.33	4.00
76	3.62	3.62	3.62	4.09	3.90	4.00	3.33
77	4.00	3.25	3.62	4.73	3.35	5.33	2.33
78	1.37	1.62	1.50	1.70	1.94	1.00	1.66
79	2.25	2.50	2.37	2.33	2.90	2.00	3.33
80	3.75	3.87	3.81	4.30	4.17	5.00	4.00
81	2.75	2.87	2.81	3.30	3.37	3.33	3.33
82	2.12	1.87	2.00	2.57	2.48	2.66	3.00
83	2.00	2.00	2.00	2.14	2.31	1.66	2.33
84	2.87	2.75	2.81	3.10	3.21	3.00	3.66
85	3.75	4.00	3.87	4.37	4.20	5.00	3.66
86	2.12	2.12	2.12	2.57	2.83	2.33	3.33
87	4.12	4.50	4.31	4.50	4.62	4.66	4.00
88	2.37	2.12	2.25	3.04	2.60	3.33	2.33
89	2.00	2.12	2.06	2.27	2.41	2.00	2.33
90	3.00	3.00	3.00	3.38	3.40	3.66	3.66
91	2.12	2.25	2.18	2.58	2.76	2.66	3.00
92	3.37	3.25	3.31	4.14	3.79	5.33	4.33
93	3.62	3.87	3.75	4.46	4.34	5.33	4.33
94	3.75	3.37	3.56	4.67	4.12	5.66	4.66
95	2.12	2.12	2.12	2.69	2.64	2.66	2.66
96	2.37	2.37	2.37	2.79	2.69	3.00	2.66
97	2.87	2.87	2.87	2.98	3.33	2.33	3.66
98	2.75	2.62	2.68	3.27	3.21	3.66	3.66
99	2.75	2.62	2.68	3.38	3.20	4.00	3.66
100	2.62	2.62	2.62	3.21	3.01	3.66	3.00
101	3.00	3.12	3.06	3.65	3.54	4.33	3.66
102	3.37	3.50	3.43	3.60	3.77	3.00	3.33

Fortsetzung Tabelle 12:

Vp	NK 1	NK 2	NK12	GN 1	GN 2	KN 1	KN 2
103	2.50	2.62	2.56	2.96	3.25	3.00	3.66
104	3.87	3.75	3.81	4.26	4.20	4.33	4.33
105	3.50	3.37	3.43	4.20	4.01	4.66	4.33
106	2.62	2.62	2.62	3.29	3.18	3.66	3.33
107	3.37	3.25	3.31	3.84	3.56	4.00	3.33
108	3.50	3.25	3.37	3.89	3.50	3.66	3.00
109	2.00	1.62	1.81	2.26	2.04	1.66	2.00
110	2.50	2.37	2.43	2.84	2.64	3.00	2.66
111	2.25	1.75	2.00	2.44	2.24	2.00	2.66
112	2.62	2.62	2.62	3.22	2.84	4.00	2.66
113	3.00	3.00	3.00	3.29	3.34	3.00	3.33
114	2.75	2.37	2.56	3.48	2.94	4.33	3.33
115	3.50	3.12	3.31	3.86	3.26	3.66	2.66
116	2.62	3.00	2.81	3.07	3.47	3.00	3.66
117	2.12	2.25	2.18	2.37	2.58	2.00	2.33
118	3.12	3.12	3.12	3.94	3.59	5.33	4.00
119	2.75	2.87	2.81	3.26	3.23	3.33	3.00
120	2.25	2.00	2.12	2.38	2.38	1.66	2.33
121	3.87	3.75	3.81	4.58	4.21	5.66	4.66
122	2.87	2.37	2.62	3.14	2.83	2.33	2.66
123	2.62	2.25	2.43	2.74	2.53	2.00	2.33
124	2.87	2.75	2.81	2.97	2.68	2.33	1.66
125	2.87	2.87	2.87	3.17	3.24	3.00	3.33
126	3.12	3.00	3.06	4.10	3.65	5.66	4.33
127	1.62	1.62	1.62	1.82	1.87	1.00	1.33
128	3.12	2.75	2.93	3.34	2.98	3.33	3.00
129	3.62	3.50	3.56	4.30	3.63	5.00	3.00
130	2.87	3.00	2.93	3.84	3.69	5.00	4.33
131	3.50	3.00	3.25	3.89	3.60	3.66	4.00
132	3.62	3.00	3.31	4.08	3.50	4.00	3.66
133	2.75	3.12	2.93	3.91	3.77	5.66	4.33
134	2.37	2.12	2.25	2.81	2.79	2.66	3.33
135	3.00	3.12	3.06	3.43	3.61	3.66	4.00
136	2.37	2.25	2.31	2.58	2.78	2.00	3.00
137	4.25	4.37	4.31	4.68	4.61	5.00	4.33
138	3.87	3.87	3.87	4.33	4.32	4.00	4.00
139	2.50	2.50	2.50	2.93	3.02	2.66	3.00
140	1.62	1.87	1.75	2.08	2.72	2.33	4.00
141	3.75	3.75	3.75	4.28	4.33	5.00	5.00
142	2.75	3.00	2.87	3.20	3.75	3.00	4.33
143	4.12	3.75	3.93	4.74	4.34	5.33	4.66
144	2.75	2.62	2.68	3.06	3.30	2.33	3.66
145	2.25	2.37	2.31	2.46	2.84	2.00	3.00
146	2.12	2.37	2.25	2.13	2.71	1.33	2.66
147	2.87	2.87	2.87	3.19	3.15	3.33	3.00
148	3.50	3.00	3.25	4.01	3.55	4.33	4.00
149	1.87	1.87	1.87	2.18	2.23	1.66	2.00
150	3.62	3.37	3.50	4.28	4.02	5.33	5.00
151	4.37	4.37	4.37	5.02	4.82	5.33	4.66
152	2.00	2.00	2.00	2.45	2.80	2.00	3.33
153	4.37	4.25	4.31	4.85	4.61	5.33	4.66
154	3.75	3.75	3.75	4.35	4.32	4.33	4.33

Fortsetzung Tabelle 12:

Vp	NK 1	NK 2	NK12	GN 1	GN 2	KN 1	KN 2
155	2.12	2.00	2.06	2.30	2.39	1.33	2.00
156	2.75	3.12	2.93	2.93	3.76	2.66	4.66
157	2.87	3.50	3.18	3.63	3.97	4.66	4.00
158	3.62	3.62	3.62	3.78	3.99	3.00	3.66
159	3.37	2.87	3.12	3.52	3.36	2.66	3.33
160	3.00	3.12	3.06	3.59	3.48	4.00	3.33
161	3.75	3.37	3.56	4.05	3.95	3.66	4.33
162	1.87	1.62	1.75	2.00	1.81	1.00	1.00
163	1.75	1.75	1.75	2.00	1.98	2.00	2.00
164	2.50	2.62	2.56	2.90	3.08	3.33	3.66
165	3.00	2.75	2.87	3.53	3.18	4.00	3.33
166	1.87	1.62	1.75	2.23	2.19	2.00	2.66
167	3.37	3.00	3.18	3.74	3.29	4.00	3.33
168	2.12	2.00	2.06	3.06	2.66	4.33	3.33
169	3.12	3.25	3.18	3.73	3.51	4.33	3.33
170	1.37	1.50	1.43	1.61	1.77	1.00	1.33
171	2.87	2.62	2.75	3.41	2.79	4.33	2.66
172	2.87	2.75	2.81	3.03	2.89	2.33	2.33
173	1.87	1.87	1.87	2.09	2.26	1.66	2.33
174	3.00	3.00	3.00	3.64	3.29	4.66	3.33
175	3.00	3.25	3.12	3.96	3.84	5.33	4.33
176	2.00	2.25	2.12	2.63	2.89	3.00	3.33
177	2.87	2.75	2.81	3.33	3.26	3.66	3.66
178	2.00	1.75	1.87	2.21	2.20	1.66	2.33
179	2.25	2.25	2.25	2.83	2.83	3.66	3.66
180	2.87	2.75	2.81	3.25	3.16	3.33	3.33
181	3.00	3.00	3.00	3.56	3.56	4.00	4.00
182	1.75	1.75	1.75	2.32	2.41	2.66	3.00
183	2.75	2.75	2.75	2.93	3.12	2.66	3.33
184	2.50	2.50	2.50	2.90	2.73	3.00	2.33
185	2.37	2.25	2.31	2.62	2.79	2.33	3.33

Abkürzungen

NK 1: Nicht-Kursfächer 1. Halbjahr
 NK 2: Nicht-Kursfächer 2. Halbjahr
 NK12: Nicht-Kursfächer 1.+2. Halbjahr
 GN 1: Gesamtnote 1. Halbjahr
 GN 2: Gesamtnote 2. Halbjahr
 KN 1: Kursfächer-Note 1. Halbjahr
 KN 2: Kursfächer-Note 2. Halbjahr

4.1.4. Diskussion

Stellt man zum Vergleich die von uns erhaltene FAT-Rohwerteverteilung (Abbildung 1) der für die 5. Klassen der Eichstichprobe gewonnenen Verteilung (Abbildung 2) gegenüber, so ist eine recht gute Übereinstimmung festzustellen. Es deutet sich für beide Stichproben eine rechtsseitige Asymmetrie der Verteilungen an, die, wie zu erwarten, für die Kurven der Real- und Gymnasialschüler (siehe Abbildung 2) etwas deutlicher ausfällt. Die für unsere Rohwerteverteilung mit den χ^2 -Anpassungstest durchgeführte Normalitätsprüfung hat ergeben, daß die Normalitätsannahme (bei einer Überschreitungswahrscheinlichkeit von 6,3 %) beizubehalten ist. Damit kann wohl für unsere Stichprobe - trotz der leichten Rechtsgipfligkeit der Verteilung - die Leistungsdifferenzierung des FAT als hinreichend angenommen werden. Die durchschnittlich erreichte Testleistung liegt für die Eichstichprobe etwas höher ($\bar{X} = 57,6$ Rohwerte¹⁾) als für unsere Stichprobe ($\bar{X} = 55,2$ Rohwerte).

Betrachtet man die für die drei AN-Verfahren erhaltenen Häufigkeitsverteilungen der ZD-Werte (Abbildung 3, 4 und 5), so fällt zunächst die Form der Verteilung auf, die bei allen drei Verfahren in sehr ähnlicher Weise wiederzufinden ist. Die Verteilung ist durchgängig asymmetrisch und fällt nach rechts steil, nach links dagegen recht flach ab. Die Kurvenmaxima (hier jeweils die Klasse mit der größten Häufigkeit)

1) Da im Beiheft zum FAT 4-6 die im Mittel erreichten Testleistungen nur getrennt nach Geschlecht und Schulart angegeben sind (BELSER und RAATZ 1973, S. 13), wurde der mitgeteilte Gesamtmittelwert als gewogenes Mittel aus den entsprechenden Angaben ermittelt.

liegen in allen drei Fällen im Bereich von ZD-Werten nahe Null (wobei ein individueller ZD-Wert von Null besagt, daß im Durchschnitt die erreichte und die anschließend vorgenommene Leistung einander entsprechen). Hält man jeweils die Bereiche positiver und negativer ZD-Werte gegeneinander, so streuen die negativen Werte deutlich weiter als die positiven Werte. Möglicherweise bedeutet dies, daß, in bezug auf die für das Zustandekommen der einzelnen ZDen konstituierenden Bedingungen, die Vpn mit positiven ZD-Werten eine in sich homogenere Gruppe bilden als die Vpn mit negativen ZD-Werten. - Qualitative Befunde aus einer phänomenologisch orientierten Studie von HECKHAUSEN (HECKHAUSEN 1955) wären geeignet, diese Vermutung zu stützen. Nach HECKHAUSEN stellen sich in bezug auf das zugrundeliegende Motivationsgeschehen positive Zieldiskrepanzen als relativ eindeutig dar, negative Zieldiskrepanzen dagegen als recht kompliziert und mehrdeutig.

Geht man von den über alle Vpn gemittelten ZD-Werten für die drei AN-Verfahren ($\overline{ZD}_{ZA} = -0,92$; $\overline{ZD}_{ZS} = -1,07$; $\overline{ZD}_{RW} = 0,32$) aus, so unterscheiden sich diese numerisch nicht sehr stark voneinander. - Für einen Vergleich zwischen den Verfahren ist allerdings zu berücksichtigen, daß numerisch gleiche ZD-Werte aus verschiedenen AN-Verfahren nicht für psychisch Gleiches stehen müssen. Zum Beispiel ist es vorstellbar, daß einem ZD-Wert von 1 beim RW ein größeres Gewicht zukommt als beim ZA. (Hierbei könnte auch die in den drei AN-Verfahren unterschiedliche Anzahl der pro Durchgang vorgegebenen Aufgaben eine Rolle spielen.) Beim RW bedeutet jeder einzelne Ring eine Leistung, wohingegen beim ZA und beim ZS die Leistung sich nur aus der Gesamtzahl gelöster Aufgaben ergibt. Die einzelne Addition

oder Substitution ist als Leistung nicht gegeben (siehe dazu auch S. 111 f.).

Nimmt man die jeweiligen Anteile positiver und negativer ZD-Werte zum Ausgangspunkt, so weicht das RW auffällig von den beiden anderen Verfahren ab. Während beim ZA und beim ZS die negativen Werte bei weitem überwiegen (von 185 Vpn haben beim ZA 106 Vpn und beim ZS 114 Vpn negative ZD-Werte)¹⁾, ist beim RW das Gegenteil der Fall (59 negative gegenüber 126 positiven ZD-Werten)¹⁾. Eine schlüssige Erklärung für diese Unterschiede in den jeweiligen Anteilen positiver und negativer ZD-Werte, die sich auf andere Untersuchungen stützen ließe, kann nicht gegeben werden. Betrachtet man die Ergebnisse vor dem Hintergrund der Aufgabenart, genauer der Art der in den verschiedenen Verfahren zu erbringenden Leistung, so kann man sagen, daß ZA und ZS als Aufgaben "schulnäher" sind und daher vermutlich zu vorsichtigeren Zielsetzungen Anlaß geben. Demgegenüber wird beim RW eher der spielerische Charakter akzentuiert, was möglicherweise die Furcht vor Mißerfolgen zurücktreten läßt und sich in risikofreudigeren Zielsetzungen niederschlägt. Eine weitere Deutungsmöglichkeit kann darin gesehen werden, daß die Verfahren ZA und ZS im Gruppenversuch, das RW im Einzelversuch durchgeführt wurden. Es wurde zwar versucht, in Hinsicht auf das AN wesentliche Unterschiede zwischen der Situation des Gruppenversuchs und der des Einzelversuchs - soweit möglich - aufzufangen, doch schien es uns zu unnatürlich und zu wenig glaubwürdig, im Einzelversuch aus Gründen der Vergleichbarkeit die jeweils

1) Die Null-Werte wurden den positiven Werten zugeschlagen.

vorgenommenen und die erreichten Leistungen ebenfalls von der Vp notieren zu lassen (siehe dazu S. 52 und S.63f.). (Wie wäre dann die Anwesenheit eines Vls noch plausibel zu machen?) Es bleibt also der Unterschied, daß die Leistungsvornahme im Einzelversuch (RW) quasi öffentlich mitgeteilt, im Gruppenversuch (ZA, ZS) eher privat notiert wird. Obgleich der Vl im Einzelversuch angehalten war, sich jeglicher Kommentare zu enthalten (siehe dazu S. 75), könnte die Einzelsituation zu einem vermehrten Ansporn und zu relativ höheren Leistungsvornahmen geführt haben. Festzuhalten bleibt, daß AN-Gruppenversuche und AN-Einzelversuche deutlich unterscheidbare situative Bedingungen aufweisen; und man kann nicht ohne weiteres davon ausgehen, daß die dabei ablaufenden psychischen Prozesse einander entsprechen. - Auf mögliche Konsequenzen aus den hier diskutierten Unterschieden (in den Anteilen positiver und negativer ZD-Werte) für die Generalität des ANs werden wir an entsprechender Stelle (siehe dazu unter Punkt 4.2.1.) noch eingehen.

Für die innerhalb eines Versuchs vorgenommenen und erreichten Leistungen wurden zunächst individuelle Mittelwerte ($\overline{AN}$, $\overline{LN}$; siehe Tabelle 6) und dann Gesamtmittelwerte, getrennt für die drei AN-Verfahren (Tabelle 6), berechnet. - Um möglichst günstige Voraussetzungen für das Zustandekommen echter AN-Bildungen und für eine hohe Zuverlässigkeit zu schaffen, sollten die AN-Aufgaben eine für die Vp angemessene mittlere Schwierigkeit besitzen. Die dieser Forderung zugrundeliegenden Überlegungen und die dazu getroffenen methodischen Maßnahmen wurden bereits dargestellt (siehe S. 42, Punkt 4, S. 49 ff., S. 65 und S. 73). - Die Ergebnisse lassen erkennen, daß es sowohl in den beiden Gruppenversuchen

wie im Einzelversuch in befriedigender Weise gelungen ist, eine mittlere Schwierigkeit vorzugeben. Die Gesamtmittelwerte (für ZA: $\bar{X}_{AN} = 12,70$, $\bar{X}_{LN} = 13,96$; für ZS: $\bar{X}_{AN} = 10,96$, $\bar{X}_{LN} = 12,52$; für RW: $\bar{X}_{AN} = 8,08$, $\bar{X}_{LN} = 7,88$) zeigen eine recht gute Annäherung an die jeweils halbe Punktzahl der pro Durchgang maximal möglichen Punktzahl ($\frac{LN_{max}}{2}$: 14,5 Aufgaben (für ZA), 12 Aufgaben (für ZS), 8 Aufgaben (für RW)).

Die im Durchschnitt erreichten ($\overline{LN}$) und die anschließend vorgenommenen ($\overline{AN}$) Leistungen wurden in ihrem Verlauf über die Versuchsdurchgänge dargestellt (Abbildung 6, 7 und 8); die zugehörigen Werte sind in Tabelle 1 des Anhangs mitgeteilt. Betrachtet man zunächst nur die Leistungsverläufe ($\overline{LN}$ -Kurven), so ist bei ZA und ZS ein deutlicher Leistungsanstieg festzustellen, wobei die Leistung beim ZS fast stetig, beim ZA mit erheblichen Schwankungen ansteigt. Der durchschnittliche Leistungszuwachs vom ersten zum letzten Durchgang beträgt bei beiden Verfahren ungefähr 5 Punkte (das heißt 5 Aufgaben). Beim RW dagegen verläuft die Leistungskurve nahezu parallel zur Abszisse. Die im Durchschnitt erreichte Endleistung liegt nur um knapp einen Punkt höher als die Anfangsleistung. Die Leistung (Anzahl der Treffer) steht also schon nach dem ersten Durchgang - zumindest im Mittel - fest.

Wenn wir die jeweils zugehörigen Durchschnittsverläufe der AN-Setzungen ($\overline{AN}$ -Kurven) in den Vergleich zwischen den Verfahren mit einbeziehen, dann zeigt sich, daß das AN beim RW merklich "besser" läuft als bei den beiden anderen Verfahren. Beim ZA und beim ZS deutet sich zum Ende des Versuchs hin - statt einer Verringerung - eher noch eine Vergrößerung der Zieldiskrepanzen an.

Demgegenüber passen sich beim RW die Leistungsvornahmen viel enger den jeweils vorausgehenden Leistungen an.

Diese Ergebnisse zum Leistungsverlauf und zum Zusammenspiel von Leistung und Leistungsvornahme werden unter dem Gesichtspunkt der Generalität des ANs noch ausführlicher erörtert (siehe unter Punkt 4.2.1.).

Die Schulleistung wurde über die Zensuren im Zwischen- (1. Halbjahr) und im Jahreszeugnis (2. Halbjahr) des 5. Schuljahres erfaßt (Tabelle 6). - Die für diese Entscheidung maßgeblichen Gründe wurden bereits genannt (siehe S. 59 f.) - Die Häufigkeiten der einzelnen Notestufen pro Fach, die getrennt für das 1. und 2. Halbjahr dargestellt sind (Tabelle 11), sind fast durchgängig eingipflig, allerdings nur zum geringeren Teil annähernd symmetrisch verteilt. Die deutliche Linksgipfligkeit, die sich für die Fächer Religion, Musik, Kunst, Werken/Handarbeit und Sport ergibt, muß wohl als ein Hinweis auf ihren relativ niederen Rang im Fächerkanon verstanden werden.

Die möglichen Fehlerquellen, mit denen die Noten als Urteile des Lehrers belastet sind, wurden in der Literatur bereits ausführlich dargestellt und diskutiert und auch in dieser Arbeit schon angesprochen (vgl. S. 59) so daß wir uns hier darauf beschränken können, auf die daraus erwachsenden Probleme (die in der Konsequenz auch die Interpretierbarkeit der Daten betreffen) hinzuweisen.

4.2. Zur Generalität des Anspruchsniveaus

Zur Beantwortung der Frage nach der Generalität des ANs (das heißt der intraindividuellen Konsistenz über verschiedene Leistungen) wird der Grad des Zusammenhangs zwischen den ZD-Werten (siehe dazu Tabelle 6) für jeweils zwei der drei AN-Verfahren (ZA, ZS und RW) bestimmt. Um den Vergleich mit anderen einschlägigen Untersuchungen zu ermöglichen, werden SPEARMANSche Rangkorrelationen berechnet.

Tabelle 13: Korrelationen (r_s) der ZD-Werte für die Verfahren ZA/ZS, ZA/RW und ZS/RW (N=185)

ZD \ ZD	ZA	ZS	RW
ZA	—	.6925 ⁺	.5876 ⁺
ZS		—	.5889 ⁺
RW			—

Für einseitige Fragestellung:

+: $P < 0,001$

4.2.1. Diskussion

Die Höhe der Interkorrelationen der ZD-Werte aus den verschiedenen AN-Verfahren (Tabelle 13) weist in allen Fällen auf einen mittleren Grad des Zusammenhangs hin, der statistisch gesichert ist. Die Ergebnisse (das heißt die Höhe der Koeffizienten) können überdies als recht stabil angesehen werden, da sie auf eine relativ breite Datenbasis (N=185) gestützt sind und der Stichprobenkorrelationskoeffizient eine konsistente Punkt-

schätzung von ρ darstellt. Die gestellte Frage nach der Generalität kann wohl dahingehend beantwortet werden, daß das AN (erfaßt über die Zieldiskrepanz) sich über verschiedene Leistungen (bzw. Aufgaben) hinweg als relativ konsistent darstellt. Die Ergebnisse sprechen dafür, das AN nicht als ein leistungsspezifisches, sondern eher als ein persönlichkeitspezifisches Merkmal aufzufassen.

Je nachdem, welchen Standpunkt man bezieht, lassen sich die gefundenen Korrelationen als relativ hoch oder auch als relativ niedrig bezeichnen. Als relativ hoch kann man die Korrelationen unter dem Gesichtspunkt ansehen, daß eigentlich sehr heterogene Leistungen - motorische Fertigkeiten, kognitive Fähigkeiten - miteinander verglichen werden. Das ZD-Maß scheint, unter diesem Blickwinkel betrachtet, geeignet zu sein, Motivationsvorgänge über die verschiedenen Leistungen hinweg - aber dennoch auf diese bezogen - abzubilden. Als relativ niedrig stellen sich die Korrelationen dar, wenn man sich die Sichtweise der Testtheorie zu eigen macht. Man kann kaum davon ausgehen, daß die drei AN-Verfahren "Paralleltests" sind, die das gleiche prüfen. Äußerlich gesehen wurden alle drei Verfahren unter möglichst vergleichbaren Bedingungen (formaler Ablauf des Versuchs, Frage nach dem AN, Urteil über erlebten Erfolg/Mißerfolg und Anzahl der Versuchsdurchgänge) durchgeführt. Durchführungs- und auch Auswertungsobjektivität, um in der Sprache der Testtheorie zu bleiben, können sicher als gewährleistet gelten. Die Gründe dafür, daß nur mittelhohe Zusammenhänge gefunden werden, müssen demnach in den Verfahren selber gesucht werden. Hierauf werden wir im folgenden ausführlicher eingehen.

Vergleicht man die Höhe der erhaltenen Koeffizienten, so fällt auf, daß die Verfahren ZA und ZS untereinander etwas höher korrelieren als jedes der beiden Verfahren mit RW. Als eine Erklärung für dieses Ergebnis bietet sich zunächst die Art der jeweils verlangten Leistung an. Während es beim RW vor allem auf motorische Geschicklichkeit ankommt, geht es bei den beiden anderen Verfahren eher um "Kopf-Leistungen". Es ergeben sich weitere Erklärungsmöglichkeiten, wenn man andere Ergebnisse aus den AN-Versuchen heranzieht.

Ein Vergleich der jeweiligen Anteile positiver und negativer ZD-Werte hatte ergeben, daß beim ZA und beim ZS die negativen ZDen deutlich überwiegen, während beim RW die positiven ZDen den weit größeren Anteil ausmachen (vgl. S.104). Diese Niveaueverschiebungen lassen die Interkorrelationen nur dann sinken, wenn sich zugleich auch die relative Position der Vp verschiebt, das heißt sich strukturelle Änderungen vollziehen. Dies scheint, wie die Ergebnisse zeigen, in gewissem Umfang zuzutreffen.

Zwischen dem RW und den beiden anderen Verfahren hatten sich weitere Unterschiede herausgestellt, die in die Generalitätsbetrachtung einbezogen werden sollen. Diese Unterschiede betreffen die Leistungsverläufe ($\overline{LN}$ -Kurven) und das Zusammenspiel der vorgenommenen und erreichten Leistungen ($\overline{AN}$ - und $\overline{LN}$ -Kurve) (siehe auch S.106 f.). Beziehen wir die entsprechenden Ergebnisse aufeinander, so bietet sich folgende Interpretation an: Nachdem beim RW (im Durchschnitt) fast kein Übungsgewinn erkennbar wird, ist hier wahrscheinlich der Spielraum, durch willentliche Anstrengung eine Leistungsverbesserung zu erzielen, größer zu veranschlagen

als beim ZA und beim ZS. Gerade dieses Sich-selbst-aktivieren-Können scheint für einen AN-Versuch ideale Voraussetzungen zu bieten und ist vermutlich noch wichtiger als die Möglichkeit, durch Übung zu einer Leistungssteigerung zu gelangen (vgl. dazu S. 43 Punkt 7. und 8.). Im Fall des RWs muß dem in der Leistungsvornahme antizipierten Ergebnis (das heißt der nächsten Leistung) eine genaue Einschätzung der eigenen Möglichkeiten zugrundeliegen. (Dabei steht es außer Frage, daß dieser Prozeß der AN-Setzung überlagert wird, z. B. durch Bescheidenheit, wenn die Vp ihr gesetztes AN mitteilen soll.) Beim ZA und beim ZS treten wohl zu viele Leistungszuwächse auf, als daß es der Vp ohne weiteres möglich wäre, ihre Leistungen und Leistungsvornahmen aufeinander zu beziehen. Man wird eben schon allein durch Übung besser. - Folgt man dieser Interpretation, dann dürften aber 11 Versuchsdurchgänge kaum ausreichen, um die Dynamik des ANs unter Lernbedingungen (wo es zusätzlich um die Einschätzung der eigenen Fähigkeit zur Leistungssteigerung geht) untersuchen zu können.

Vor dem Hintergrund der Generalität bietet sich eine weitere Interpretationsmöglichkeit für die Unterschiede in den $\overline{AN}$ - und $\overline{LN}$ -Verläufen an. Vergleicht man die drei AN-Verfahren unter dem Aspekt der jeweils gegebenen Leistungssituation, dann ist folgendes festzuhalten: Beim RW stellt sich die Leistungssituation überschaubarer, strukturierter dar. Der Verlauf der Leistung während eines Durchgangs ist in jeder Phase gegenwärtig. Die Rückmeldung über die erbrachte Leistung ist sofort und unmittelbar anschaulich gegeben (Treffer oder Fehlwurf). Die Leistung kann sich für die Vp zum Beispiel in Treffer, beinahe Treffer und Fehl-

würfe (ungeschickter Wurf oder "Pech gehabt") gliedern. Beim ZA und beim ZS dagegen gestaltet sich die Leistungssituation weniger anschaulich und übersichtlich. Fast könnte man von einem "Ankämpfen gegen eine imaginäre Zeitvorgabe" sprechen. Die Rückmeldung über die Leistung ist nicht sofort (das heißt nach jeder einzelnen Aufgabe eines Durchgangs) und direkt anschaulich gegeben, sondern erfolgt erst am Ende eines Durchgangs. Der Leistungsverlauf innerhalb eines Durchgangs ist daher nicht präsent. - In der Literatur finden sich zu diesem Gesichtspunkt der Anschaulichkeit der Leistungssituation keine Hinweise. Doch kann man auf Grund der hier vorliegenden Ergebnisse wohl annehmen, daß diese Anschaulichkeit für die AN-Setzung bzw. für das Zustandekommen der Zieldiskrepanzen eine wichtige Rolle spielt. - Diese Frage sollte in einer eigenen Untersuchung weiterverfolgt werden.

Die zur Generalität des ANs dargestellten Untersuchungen (FRANK 1935; GOULD 1938, 1939) wurden beide an studentischen Stichproben durchgeführt; auch noch in anderer Hinsicht liegen unterschiedliche Versuchsbedingungen vor, so daß eine direkt vergleichende Gegenüberstellung nicht zulässig ist. Die bei FRANK (1935) mitgeteilten Rangkoeffizienten differieren mit Werten zwischen $-.08$ und $.70$ beträchtlich, wobei erstaunlicherweise die auf verschiedene AN-Verfahren (und zugleich verschiedene Stichproben) bezogenen Interkorrelationen weniger stark voneinander abweichen als die Koeffizienten, die sich jeweils auf eine erste und eine nach einer Woche wiederholte Messung an denselben Vpn beziehen. Bedenkt man, daß diese sehr unterschiedlich hohen Koeffizienten jeweils nur auf den Daten von 12 Vpn basieren, dann muß man wohl sagen, daß die Unter-

suchung von FRANK die Frage der Generalität des ANs noch weitgehend offen läßt. GOULD (1938, 1939) findet Interkorrelationen, die zwischen .04 und .44 streuen und im Mittel einen Koeffizienten von .29 ergeben. GOULD, die ihre Untersuchung an 82 männlichen Studenten durchführte, sieht in ihren Ergebnissen keine Bestätigung für die von FRANK angenommene Generalität des ANs.

4.3. Zur Stabilität des Anspruchsniveaus

Zur Schätzung der in Frage stehenden Stabilität des ANs werden die AN-Daten der ersten Versuchsdurchführung (Hauptversuche I) zu den entsprechenden Daten der wiederholten Durchführung (Hauptversuche II) in Beziehung gesetzt. Der Grad des Zusammenhangs zwischen den Ergebnissen der 1. und 2. Messung wird zunächst über die ZD-Werte getrennt für die drei AN-Verfahren (ZA, ZS und RW) bestimmt. Zusätzlich wird eine Retestkorrelation berechnet, die auf einem AN-Gesamtmaß über alle drei Verfahren basiert.

Ausgangspunkt für die Berechnung dieses Gesamtmaßes war die Überlegung, daß ein auf die jeweiligen Ergebnisse in allen drei Verfahren gestütztes Maß eine größere Stabilität (nicht im Sinne zeitlicher Stabilität zu verstehen) erwarten lasse als ein lediglich auf ein einzelnes Verfahren bezogenes Maß. Die Höhe der Interkorrelationen der AN-Verfahren (siehe Tabelle 13) scheint geeignet, diese Maßnahme zu rechtfertigen. - Das AN-Gesamtmaß ($\overline{ZD}_{rel.}$) wurde folgendermaßen berechnet: Mit dem Ziel, die individuellen ZD-Werte über die drei AN-Verfahren hinweg vergleichbar zu machen, wurden diese zu-

erst auf die pro Durchgang vorgegebene Aufgabenzahl des jeweiligen Verfahrens (ZA: 29 Aufgaben; ZS: 24 Aufgaben; RW: 16 Aufgaben) relativiert. Außerdem wurden die ZD-Werte mit dem Faktor 100 multipliziert, um nicht zu kleine Werte zu erhalten und um Rundungsfehler auszuschließen.

$$ZD_{\text{rel.}} = \frac{ZD \cdot 100}{\text{Anzahl der Aufgaben pro Durchgang}}$$

Die so erhaltenen $ZD_{\text{rel.}}$ -Werte wurden dann für jede Vp über alle drei AN-Verfahren gemittelt.

Tabelle 14: ZD-Werte (für ZA, ZS, RW) und $\overline{ZD}_{rel.}$ -Werte der 1. und der 2. Messung (N = 60)

Nr.	Vp	ZA		ZS		RW		$\overline{ZD}_{rel.}$	
		ZD _I	ZD _{II}	ZD _I	ZD _{II}	ZD _I	ZD _{II}	I	II
1	3	-0,4	0,6	-0,6	-1,2	0,4	1,0	- 0,46	1,11
2	6	1,0	0,5	1,9	0,4	1,3	0,7	6,50	2,59
3	8	-0,7	-1,1	-0,8	-0,9	0,0	-1,4	- 1,91	- 5,43
4	10	1,2	1,2	1,9	2,1	1,8	1,4	7,77	7,21
5	11	-0,6	-0,1	-0,5	-0,5	1,0	0,7	0,70	0,65
6	13	-0,4	-3,4	-0,3	-1,6	0,3	0,2	- 0,25	- 5,71
7	20	0,7	0,6	-1,2	1,2	1,3	-0,3	1,84	1,73
8	21	-3,6	-1,2	-5,7	-3,0	-1,6	0,0	-15,39	- 5,55
9	25	-0,3	-0,5	-0,9	-1,9	-0,8	-0,9	- 3,26	- 5,09
10	29	2,6	-0,8	-0,1	-0,1	1,2	2,5	5,35	4,15
11	30	2,2	2,4	3,4	0,4	1,2	0,2	9,75	3,73
12	32	0,4	-1,8	-1,4	-3,7	1,3	-1,2	1,22	- 9,71
13	34	0,3	-2,4	-2,3	-4,1	0,9	-0,3	- 0,98	- 9,08
14	35	-1,1	3,4	-0,1	1,8	1,3	2,3	1,30	11,20
15	42	1,5	-0,8	1,8	-0,1	1,3	0,4	6,93	- 0,23
16	44	-1,7	-1,4	-4,2	-1,0	-1,1	-0,6	-10,08	- 4,25
17	47	-2,3	-1,6	-0,4	1,5	1,5	0,4	- 0,07	1,08
18	49	-1,5	-4,2	-1,8	-2,2	0,4	0,9	- 3,39	- 6,01
19	53	2,3	2,9	-2,3	3,1	3,5	2,3	6,74	12,43
20	55	-0,1	-0,2	1,1	-0,9	0,9	-0,1	3,29	- 1,69
21	57	0,7	0,9	-0,1	-0,4	3,0	0,4	6,91	1,31
22	60	-2,3	-0,4	-2,3	-1,6	0,8	1,2	- 4,17	- 0,18
23	64	-3,4	-4,5	-1,3	-3,4	-0,7	-2,9	- 7,17	-15,94
24	65	-4,7	-6,2	-2,2	-5,6	-0,8	-0,9	-10,13	-16,78
25	71	-0,6	-2,0	-1,4	-3,4	1,1	-0,6	- 0,34	- 8,27
26	74	1,2	3,5	2,6	5,2	2,8	3,3	10,82	18,12
27	76	0,2	-1,0	-0,8	-1,3	0,2	-1,6	- 0,46	- 6,29
28	85	-4,7	-2,6	-4,5	-0,2	-0,8	-1,1	-13,32	- 5,56
29	92	-0,3	1,2	0,2	-1,1	1,6	-0,7	3,27	- 1,61
30	94	1,5	2,4	1,1	3,3	1,8	2,5	7,00	12,55

Fortsetzung Tabelle 14

Nr.	Vp	ZA		ZS		RW		$\overline{ZD}_{rel. II}$	
		ZD _I	ZD _{II}	ZD _I	ZD _{II}	ZD _I	ZD _{II}	I	II
31	95	0,9	0,8	1,5	-0,9	1,4	1,1	6,03	1,96
32	101	-1,9	-5,1	-2,2	-4,7	-1,0	0,1	- 7,32	-12,18
33	103	2,6	0,7	1,9	1,3	1,7	1,4	9,17	5,53
34	104	-1,0	-1,1	-1,4	-1,1	0,0	0,1	- 3,09	- 2,58
35	111	-5,2	-6,8	-7,6	-5,8	-1,9	-0,4	-20,49	-16,70
36	114	0,8	0,5	1,4	0,1	0,4	0,2	3,70	1,13
37	116	0,9	0,2	1,2	0,7	0,9	1,2	4,57	3,70
38	119	-2,3	-1,5	0,3	-0,8	0,4	-0,1	- 1,39	- 3,04
39	124	-1,3	0,0	-0,8	0,4	0,0	2,0	- 2,60	4,72
40	125	0,2	0,4	-0,3	0,2	1,2	1,0	2,31	2,82
41	128	-0,6	0,2	0,1	0,5	0,4	0,5	0,28	1,97
42	133	-0,6	-1,6	0,2	0,5	0,1	-2,2	- 0,21	- 5,73
43	137	2,4	1,5	0,4	1,6	1,5	1,8	6,44	7,70
44	138	0,2	-0,7	0,4	-0,2	0,6	1,2	2,04	1,42
45	139	-1,1	-1,3	-2,1	-1,6	0,5	-0,3	- 3,14	- 4,34
46	140	-5,1	-4,9	-5,8	-5,0	0,5	0,3	-12,88	-11,95
47	145	-3,3	-1,8	-3,1	-1,3	-0,7	-1,1	- 9,56	- 6,17
48	151	0,5	1,3	0,5	0,4	0,4	-0,4	2,10	1,22
49	152	-3,6	-1,8	-1,8	-1,1	-0,8	0,3	- 8,30	- 2,97
50	153	-4,4	-4,2	-2,8	-3,2	-1,6	-1,3	-12,28	-11,98
51	154	-0,6	1,6	-2,0	-0,2	1,0	0,0	- 1,38	1,56
52	162	-5,0	-2,9	-2,3	-1,0	-0,6	-0,4	-10,19	- 5,56
53	164	-1,2	-0,8	0,1	-1,1	-0,2	-1,3	- 1,66	- 5,16
54	165	-1,4	-5,3	-2,3	0,0	0,2	0,7	- 4,39	- 4,63
55	167	-0,9	-1,0	0,3	0,2	1,1	-0,4	1,67	- 1,70
56	170	0,0	-1,1	1,6	0,0	1,1	0,2	4,51	- 0,85
57	177	1,3	1,8	0,8	0,1	0,9	0,6	4,48	3,46
58	179	-5,6	-1,5	-3,2	-2,1	-0,2	0,5	-11,30	- 3,60
59	182	-3,9	-2,7	-2,1	-2,3	0,0	-1,6	- 7,40	- 9,63
60	185	-3,7	-3,0	-4,7	-2,4	-0,9	-1,1	-12,65	- 9,07
$\bar{X}$		-0,93	-0,94	-0,92	-0,80	0,52	0,17	- 1,25	- 1,84
s		2,17	2,29	2,15	2,09	1,20	1,22	6,99	7,07

Die für die AN- und LN-Werte der Reteststichprobe berechneten Gesamtmittelwerte und -standardabweichungen sind in Tabelle 15 aufgeführt.

Tabelle 15: Mittelwerte und Standardabweichungen der AN- und LN-Werte, getrennt für ZA, ZS und RW (N = 60)

AN-Verfahren	AN		LN	
	$\bar{X}$	s	$\bar{X}$	s
ZA	12,08	4,10	13,38	4,03
ZS	11,28	3,19	12,58	3,13
RW	8,58	2,41	8,53	2,44

Die ermittelten Retestkorrelationen sind in Tabelle 16 zusammengestellt. Auch hier wird aus Gründen der Vergleichbarkeit mit anderen Untersuchungen zur Stabilität des ANs ein r_s -Koeffizient berechnet.

Tabelle 16: Retestkorrelationen (r_s) der ZD-Werte (für ZA, ZS, RW) und der $\overline{ZD}_{rel.}^s$ -Werte der 1. und der 2. Messung (N = 60)

	ZD_{ZA}	ZD_{ZS}	ZD_{RW}	$\overline{ZD}_{rel.}^s$
r_s	.7444 ⁺	.6677 ⁺	.5793 ⁺	.7780 ⁺

Für einseitige Fragestellung:

+: $P < 0,001$

4.3.1. Diskussion

Die Frage nach intraindividuellen Unterschieden bzw. Übereinstimmungen des ANs galt zum einen der Konsistenz über verschiedene Leistungen (Generalität), zum anderen der Konsistenz über die Zeit (Stabilität).

Betrachten wir zuerst die über alle Vpn (N = 60) gemittelten ZD-Werte der 1. und der 2. Messung (Tabelle 14), so zeigen sich nur geringfügige (bei ZA und ZS) bis leichte (bei RW) Unterschiede. Aus diesen Mittelwertsergebnissen sind allerdings noch keine zwingenden Rückschlüsse auf entsprechende intraindividuelle Gegebenheiten zu ziehen. Wesentlicheren Aufschluß über die Stabilität geben die in Tabelle 16 mitgeteilten Retestkorrelationen. Die Größenordnung der Koeffizienten spricht klar für die Stabilität des (über ZD-Werte erfaßten) ANs; zumal dann, wenn man berücksichtigt, daß das Zeitintervall zwischen beiden Messungen mit 3 1/2 Monaten recht lang bemessen war. Man darf wohl auch annehmen, daß dieser Zeitabstand genügend groß gewählt war, um einerseits Erinnerungseinflüsse (die zu einer Überschätzung der Stabilität führen würden) weitgehend ausschließen zu können und andererseits etwaige Merkmalsschwankungen zur Wirkung gelangen zu lassen.

Die Korrelation für das auf allen drei AN-Verfahren basierende Gesamtmaß ($\overline{ZD}_{rel.}$) fällt erwartungsgemäß am höchsten aus ($r_s = .7780$). - Ein Koeffizient dieser Höhe würde bereits den Anforderungen genügen, die üblicherweise an die Retestreliabilität standardisierter Tests gestellt werden (vgl. etwa LIENERT 1967, S. 309). - Bei den über die ZD-Werte, getrennt für die

einzelnen Verfahren bestimmten Korrelationen liegt der Koeffizient für ZA am höchsten ($r_s = .7444$); der Koeffizient für ZS ($r_s = .6677$) fällt dagegen nur leicht, der Koeffizient für RW ($r_s = .5793$) etwas mehr ab. Vergewärtigen wir uns, daß das AN im Durchschnitt gesehen beim RW besser zu "funktionieren" schien (das heißt sich den jeweils vorausgehenden Leistungen besser annäherte) als bei den beiden anderen Verfahren (vgl. Abbildung 6, 7, 8 und S. 106 f.), so mag dieses Ergebnis zunächst erstaunen. Als ein systematischer Faktor kommt die Verschiebung des Gesamtmittels der ZD-Werte (von der 1. zur 2. Messung) in Betracht (siehe oben und Tabelle 14). Eine insgesamt gegebene Niveaushverschiebung kann sich jedoch nur insoweit stabilitätsmindernd auswirken, als sie zugleich mit strukturellen Änderungen (von der 1. zur 2. Messung) einhergeht. Wenn dies beim RW in verstärktem Maß der Fall wäre, müßte die entsprechende Korrelation allerdings niedriger ausfallen.

Eher ist daher eine andere Erklärung zu erwägen. Dazu sind vorab einige Erläuterungen zum Maß der Zieldiskrepanz erforderlich. Die ZD ist ein gegenüber Ausreißerwerten recht empfindliches Maß. Solche Ausreißerwerte können dann entstehen, wenn die Vp in einem Durchgang eine Leistung erbringt, die deutlich - nach oben oder unten - aus dem Bereich ihrer individuellen Leistungsfähigkeit herausfällt ("Glück" oder "Pech" gehabt) und daher bei der nachfolgenden AN-Setzung unberücksichtigt bleibt. Zwar werden sich derartige Extremwerte auf alle (individuell) gemittelten Zieldiskrepanzen auswirken, aber bei insgesamt hohen ZD-Werten weniger stark durchschlagen als bei niedrigen ZD-Werten. Diese Fehlerquelle geht auch in den intraindividuellen

Vergleich mit ein, wobei der stabilitätsmindernde Effekt bei (dem Absolutbetrag nach) kleineren ZD-Werten höher zu veranschlagen ist. - Beziehen wir diese Überlegungen auf unsere Ergebnisse, so wäre zu fragen, ob die Betrachtung der $|ZD|$ -Werte zu Unterschieden zwischen den AN-Verfahren führt. Nimmt man von den $|ZD|$ -Werten (diese sind der Tabelle 14 zu entnehmen, wenn man die Vorzeichen unberücksichtigt läßt) die Mittelwerte als Anhalt, dann zeigt sich, daß die mittleren $|ZD|$ -Werte bei ZS (1. Messung: $\bar{X}_{|ZD|} = 1,74 \pm 1,55$; 2. Messung: $\bar{X}_{|ZD|} = 1,63 \pm 1,52$) und vor allem bei ZA (1. Messung: $\bar{X}_{|ZD|} = 1,78 \pm 1,54$; 2. Messung: $\bar{X}_{|ZD|} = 1,90 \pm 1,58$) höher liegen als bei RW (1. Messung: $\bar{X}_{|ZD|} = 0,98 \pm 0,71$; 2. Messung: $\bar{X}_{|ZD|} = 0,95 \pm 0,78$). - Was zunächst verwunderte, nämlich eine augenscheinlich leicht geringere Stabilität beim RW, läßt sich nun eher erklären.

Die Gesamtmittelwerte der in den einzelnen Verfahren jeweils vorgenommenen (AN) und erreichten (LN) Leistungen (Tabelle 15) zeigen an, daß es - ebenso wie in den Hauptversuchen I (vgl. S.105f.)- auch für die Versuchswiederholung in guter Annäherung gelungen ist, einen mittleren Schwierigkeitsbereich einzuführen. Damit sind zugleich - auf Grund der theoretischen Zusammenhänge zwischen Varianz und Reliabilität (vgl. S. 50 oben) - günstige Voraussetzungen für eine hohe Zuverlässigkeit gegeben.

Wenn die in Frage stehende Stabilität des ANs sich erweisen soll, so ist hierfür sinnvollerweise zumindest vorauszusetzen, daß das Zeitintervall zwischen einer ersten und einer wiederholten Messung genügend lang gewählt wird. Dieser Zeitabstand betrug bei FRANK (1935) 1 Woche, bei ROTTER (1942) 4 Wochen und bei BRUCKS (1977)

19 Wochen. ROTTER, der seine Untersuchung an Erwachsenen durchführte, fand einen Retestkoeffizienten von .32.¹⁾ Bei BRUCKS ergab sich ein Koeffizient von $r_s = .54$. Zwar deutet dieses Ergebnis, nachdem es auf einem relativ langen Zeitintervall basiert, bereits eine gewisse Merkmalskonstanz des ANs an; doch ist einschränkend zu berücksichtigen, daß die Untersuchung an 4- bis 6jährigen Kindern durchgeführt wurde (das heißt ein Altersbereich, in dem entwicklungspsychologisch gesehen die Anfänge der AN-Bildung liegen).

4.4. Zur Beziehung zwischen den Variablen Anspruchsniveau, Intelligenz und Schulleistung

Um eine Aussage über die Beziehung zwischen AN, Intelligenz und Schulleistung zu ermöglichen, werden zunächst die Zusammenhänge zwischen jeweils zwei der drei Variablen betrachtet. In einem weiteren Auswertungsschritt werden unter Konstanthaltung je einer der drei Variablen partielle Korrelationen bestimmt. Schließlich wird eine multiple Korrelation berechnet, als ein Maß des Zusammenhangs der Schulleistung mit den beiden Merkmalen AN und Intelligenz.

Die zwischen den Kennwerten der drei Variablen ermittelten Korrelationen sind in Tabelle 17 dargestellt. Es wurden KENDALL-BURR's τ -Koeffizienten berechnet (LIENERT 1973, S. 625 ff.). Die jeweils zugrundeliegenden Daten sind der Tabelle 6 (ZD-Werte für ZA, ZS, RW; FAT-Rohwerte) und der Tabelle 12 (Notendurchschnitte) zu entnehmen.

1) Es fehlen Angaben zur Art der berechneten Korrelation und zur Signifikanz.

Tabelle 17: Interkorrelationen (KENDALL-BURRS τ) der Kennwerte für die Variablen AN, Intelligenz und Schulleistung (N = 185)

	NK 1	NK 2	NK 12	GN 1	GN 2	KN 1	KN 2	ZA ¹⁾	ZS ¹⁾	RW ¹⁾	FAT (RW)
NK 1	—	+ .8393+++	+ .9265+++	+ .8531+++	+ .7646+++	+ .6911+++	+ .4300+++	- .1779++	- .0820	- .1082	+ .5920+++
NK 2		—	+ .9301	+ .8041+++	+ .8639+++	+ .6458+++	+ .5062+++	- .1734++	- .0917	- .0948	+ .5791+++
NK 12			—	+ .8505	+ .8225+++	+ .6765+++	+ .4557+++	- .1758++	- .0811	- .1004	+ .5070+++
GN 1				—	+ .7538	+ .8337+++	+ .4550+++	- .1840+	- .0964	- .1236+	+ .6296+++
GN 2					—	+ .6324+++	+ .6736+++	- .1184+	- .0198	- .0579	+ .5955+++
KN 1						—	+ .4495	- .1938++	- .0876	- .1905	+ .6503+++
KN 2							—	- .0100	+ .0276	- .0445	+ .4492+++
ZA ¹⁾								—			- .1343+
ZS ¹⁾									—		- .0391+
RW ¹⁾										—	- .1401+

1) Für die Beziehungen AN/Intelligenz und AN/Schulleistung wurde zweiseitig geprüft (für alle übrigen Korrelationen wurde einseitig geprüft)

Abkürzungen:

NK 1: Nicht-Kursfächer 1. Halbjahr
 NK 2: Nicht-Kursfächer 2. Halbjahr
 NK 12: Nicht-Kursfächer 1.+2. Halbjahr
 GN 1: Gesamtnote 1. Halbjahr
 GN 2: Gesamtnote 2. Halbjahr
 KN 1: Kursfächer-Note 1. Halbjahr
 KN 2: Kursfächer-Note 2. Halbjahr

+ : $P < 0,05$
 ++ : $P < 0,01$
 +++ : $P < 0,001$

Ergänzend zu dieser Korrelationsmatrix (Tabelle 17) soll die Beziehung zwischen AN (ZD-Werte) und Schulleistung (GN 1) etwas näher betrachtet werden. Reduziert man durch entsprechende Gruppierung die Rangreihen der ZD-Werte (getrennt für ZA, ZS, RW) und die GN 1-Werte auf je 5 Stufen (--/-/0/+ /++), so läßt sich die Beziehung beider Variablen in binordinalen Kontingenztafeln darstellen (Tabelle 18).

Tabelle 18: Kontingenztafeln der 5stufig gruppierten GN 1-Werte und ZD-Werte, getrennt für ZA, ZS und RW

AN (ZD) GN 1	ZA					Σ
	++	+	0	-	--	
++	6	3	6	8	13	36
+	7	6	6	9	9	37
0	5	13	5	10	4	37
-	8	8	10	5	6	37
--	10	10	8	6	4	38
Σ	36	40	35	38	36	185
AN (ZD) GN 1	ZS					Σ
	++	+	0	-	--	
++	9	2	7	10	8	36
+	7	5	7	8	10	37
0	6	9	12	6	4	37
-	5	10	7	6	9	37
--	12	6	7	9	4	38
Σ	39	32	40	39	35	185
AN (ZD) GN 1	RW					Σ
	++	+	0	-	--	
++	4	6	6	8	12	36
+	7	6	10	6	8	37
0	9	5	12	9	2	37
-	8	9	3	9	8	37
--	12	8	4	6	8	38
Σ	40	34	35	38	38	185

Für die Gruppe der guten (GN 1 ++) und der schlechten (GN 1 --) Schüler sind die Häufigkeitsverteilungen der ZD-Werte (jeweils für ZA, ZS und RW) in den Abbildungen 9 und 10 veranschaulicht.

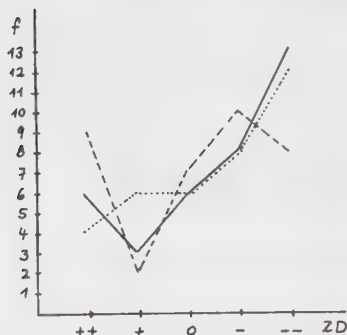


Abbildung 9:

Häufigkeitsverteilung der ZD-Werte für ZA (—), ZS (---) und RW (....) für die Gruppe der guten Schüler (GN 1 ++); N = 36

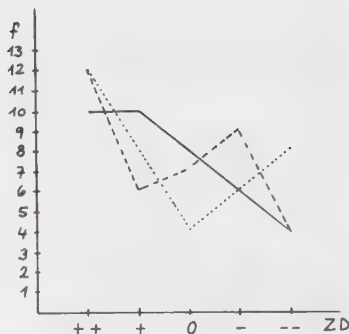


Abbildung 10:

Häufigkeitsverteilung der ZD-Werte für ZA (—), ZS (---) und RW (....) für die Gruppe der schlechten Schüler (GN 1 --); N = 38

Um festzustellen, ob bzw. inwieweit die gefundenen τ -Korrelationen (siehe Tabelle 17) durch die jeweils dritte Variable bedingt sind, werden partielle Korrelationen bestimmt. Es werden KENDALLS partielle τ -Korrelationen berechnet. Dieses partielle τ kann bei großem Stichprobenumfang aus den 3 τ -Interkorrelationen ge-

schätzt werden (LIENERT 1973, S. 658 f.). Tabelle 19 enthält die ermittelten partiellen τ -Koeffizienten, die zum Vergleich den jeweils zugehörigen τ -Koeffizienten gegenübergestellt sind.

Tabelle 19: Partielle Korrelationen (KENDALLs partielles τ) für die Zusammenhänge zwischen den Variablen AN (ZD-Werte für ZA, ZS, RW), Intelligenz (FAT-Rohwerte) und Schulleistung (GN 1); N = 185

partielles τ	τ
$\tau_{x_1 y.z} \quad -.0242$	$\tau_{x_1 y} \quad -.1343$
$\tau_{x_2 y.z} \quad .0279$	$\tau_{x_2 y} \quad -.0391$
$\tau_{x_3 y.z} \quad -.0808$	$\tau_{x_3 y} \quad -.1401$
$\tau_{x_1 z.y} \quad -.1292$	$\tau_{x_1 z} \quad -.1840$
$\tau_{x_2 z.y} \quad -.0925$	$\tau_{x_2 z} \quad -.0964$
$\tau_{x_3 z.y} \quad -.0460$	$\tau_{x_3 z} \quad -.1236$
$\tau_{yz.x_1} \quad .6210$	$\tau_{xy} \quad -.6296$
$\tau_{yz.x_2} \quad .6292$	
$\tau_{yz.x_3} \quad .6232$	

Bezeichnung der Variablen:

x_1 : ZD (ZA) y : FAT (RW)

x_2 : ZD (ZS) z : GN 1

x_3 : ZD (RW)

Sowohl für die partielle wie für die (nachfolgend berechnete) multiple τ -Korrelation ist bislang die Stichprobenverteilung des betreffenden τ -Koeffizienten nicht

bekannt, so daß derzeit noch keine Signifikanztests möglich sind (vgl. LIENERT 1973, S. 658 und 660).

Als Maß des Zusammenhangs der Schulleistung (GN 1) mit den beiden Merkmalen AN (ZD-Werte für ZA, ZS, RW) und Intelligenz (FAT-Rohwerte) wird MORANs multiple τ -Korrelation ($\tau_{x(yz)}$) errechnet (LIENERT 1973, S. 660 f.). Die erhaltenen Koeffizienten sind in Tabelle 20 mitgeteilt.

Tabelle 20: Multiple Korrelationen (MORANs multiples τ) für den Zusammenhang der Schulleistung (GN 1) mit den Variablen Intelligenz (FAT-Rohwerte) und AN (ZD-Werte; getrennt für ZA, ZS, RW); N = 185

multiples τ	
$\tau_{x(yz_1)}$	.6375
$\tau_{x(yz_2)}$	.6337
$\tau_{x(yz_3)}$	.6306

Bezeichnung der Variablen:

x: GN 1
y: FAT (RW)
z₁: ZDZA
z₂: ZDZS
z₃: ZDRW

Abschließend werden zur Übersicht die zu den Fragen

- der Beziehung zwischen den Variablen AN, Intelligenz und Schulleistung,
 - der Generalität des ANs und
 - der Stabilität des ANs
- ermittelten Korrelationen in einer schematischen Darstellung zusammengefaßt.

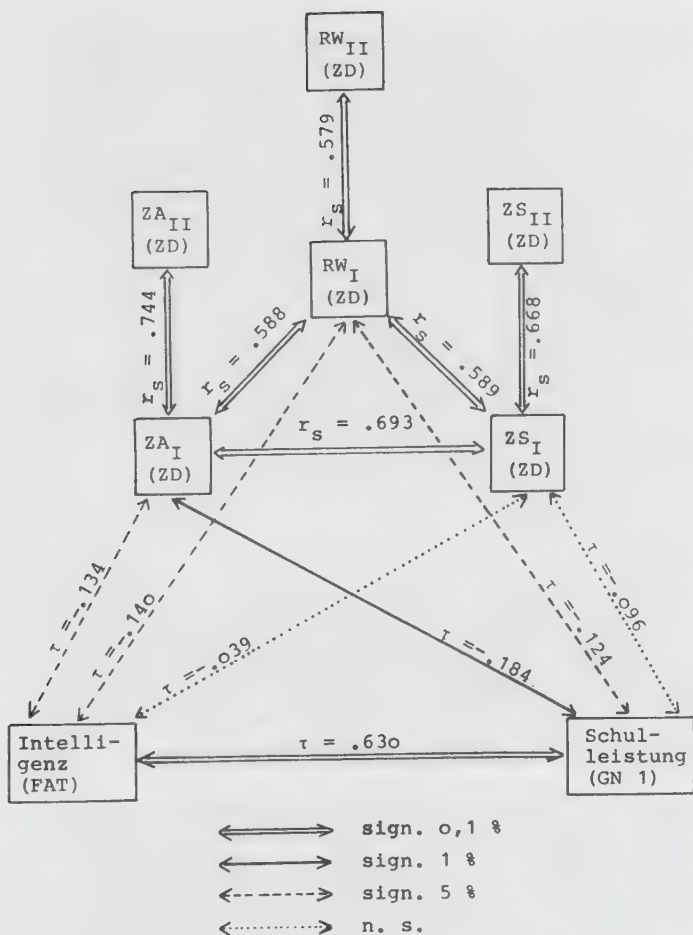


Abbildung 11: Graphische Darstellung der Korrelationen (entnommen aus den Tabellen 13, 16 und 17).

Aus Gründen der Vergleichbarkeit mit anderen einschlägigen Untersuchungen wurde für die zur Frage der Generalität und der Stabilität des ANs ermittelten Korrelationen ein r_s -Koeffizient einem τ -Koeffizienten vorgezogen. - Zur Interpretation eines τ -Koeffizienten ist auf folgendes hinzuweisen. Zwar sind alle drei gängigen Korrelationskoeffizienten (PEARSON-BRAVAIS' r , SPEARMANs r_s und KENDALLs τ) aus der gleichen allgemeinen Formel ableitbar (KENDALL 1970, S. 19 ff.), jedoch ist daraus nicht zu folgern, daß die numerischen Werte der entsprechenden Korrelationskoeffizienten gleich hoch ausfallen. Es gilt allgemein, daß sich beim gleichen Datensatz für das τ immer ein kleinerer Wert ergibt als für das r_s .

4.4.1. Diskussion

Um über die in Frage stehenden Zusammenhänge zwischen AN, Intelligenz und Schulleistung Aufschluß zu erhalten, wurden die Kennwerte dieser drei Variablen zueinander in Beziehung gesetzt. Die ermittelten Interkorrelationen sind in Tabelle 17 zusammengestellt.

Betrachten wir zuerst die Korrelation zwischen Intelligenz (FAT-Rohwerte) und Schulleistung (verschiedene Notendurchschnitte). Die Größenordnung der erhaltenen Koeffizienten (die τ -Koeffizienten bewegen sich zwischen .4492 und .6503) entspricht den in der Literatur berichteten Ergebnissen. Es werden je nach den speziellen Gegebenheiten der Untersuchung (z. B. Umfang und Zusammensetzung der Stichprobe, Art der Datenerhebung zu den beiden Variablen usw.) unterschiedliche, im

allgemeinen mittelhohe Korrelationen gefunden. Die Korrelationen mit den Notendurchschnitten des 1. Halbjahres (NK 1, GN 1, KN 1) fallen durchweg höher aus als die entsprechenden Korrelationen mit den Notendurchschnitten des 2. Halbjahres (NK 2, GN 2, KN 2). Hierbei ist die (entsprechend dem Gesamtschulkonzept) im 2. Halbjahr eingeführte Leistungsdifferenzierung in den Fächern Deutsch, Englisch und Mathematik zu berücksichtigen (vgl. S. 60). Im Hinblick auf die bei der Notenvergabe eingehenden Beurteilungsmaßstäbe bedeutet dies, daß sich für den Schüler vom 1. zum 2. Halbjahr die Referenzgruppe ändert. An die Stelle des Klassendurchschnitts im 1. Halbjahr tritt im 2. Halbjahr der Leistungsdurchschnitt des jeweiligen Kurses. Konsequenz dessen ist, daß die Korrelationen mit denjenigen Notendurchschnitten des 2. Halbjahres, die sich auf die drei Kursnoten beziehen (KN 2) bzw. diese Noten mit einschließen (GN 2), kaum mehr als interpretierbar zu bezeichnen sind. Dies gilt entsprechend für die Korrelationen zwischen den betreffenden Notendurchschnitten des 1. und 2. Halbjahres (GN 1/GN 2; KN 1/KN 2). Dagegen kann die relativ hohe Korrelation zwischen den Notendurchschnitten der Nicht-Kursfächer im 1. und 2. Halbjahr ($NK\ 1/NK\ 2 : r = .8393$) eher im Sinne eines Stabilitätsmaßes aufgefaßt werden; wobei allerdings auch hier die Frage bleibt, inwieweit man von einer Stabilität des Beurteilten (das heißt der Schülerleistung) reden kann bzw. von einer Stabilität des Beurteilten reden muß.

Unter den zwischen der FAT-Testleistung und den verschiedenen Notendurchschnitten ermittelten Korrelationen fällt diejenige mit der Durchschnittsnote der Fächer Deutsch, Englisch und Mathematik im 1. Halbjahr

(KN 1) am höchsten aus ($r = .6503$). Demgegenüber ergibt sich mit der Durchschnittsnote der übrigen Fächer¹⁾ im 1. Halbjahr (NK 1) eine etwas geringere Korrelation ($r = .5920$). Der FAT erfaßt im wesentlichen die Fähigkeit, Gesetzmäßigkeiten zu erkennen und anzuwenden. Diese Fähigkeit dürfte in den Fächern Deutsch, Englisch und Mathematik eine größere Rolle spielen, als dies für die Mehrzahl der übrigen Fächer (und zwar Religion, Musik, Kunsterziehung, Handarbeit/Werken und Sport) anzunehmen ist.

Die aus Gültigkeitsuntersuchungen zum FAT (siehe Beiheft zum FAT, BELSER und RAATZ 1973, S. 10 f.) stammenden Korrelationen zwischen FAT-Testleistung und Schulleistung (Zeugnisnoten in Deutsch und Mathematik) liegen mit Koeffizienten zwischen .12 und .40²⁾ (vgl. auch Tabelle 5 S. 58) deutlich niedriger als die in der vorliegenden Arbeit erhaltenen Korrelationen. Für diese Unterschiede wäre folgende Interpretation denkbar: Die bei BELSER und RAATZ mitgeteilten Korrelationen beziehen sich jeweils nur auf Schüler einer Schulart (Hauptschule, Realschule, Gymnasium). Folglich werden die dort erhobenen FAT-Rohwerte innerhalb der drei Schulstichproben wahrscheinlich weniger weit streuen, als dies bei einer Gesamtschulstichprobe mit einem vergleichbaren Stichprobenumfang der Fall wäre. Bedenkt man weiter, daß eine geringere Streuung der Rohwerte die Korrelation sinken läßt, so läge darin eine mögliche Erklärung für die relativ niedrigen Korrelationen bei BELSER und RAATZ. - Diese Über-

1) Dies sind die Fächer Religion, Erdkunde, Physik/Chemie, Biologie, Musik, Kunsterziehung, Handarbeit bzw. Werken und Sport.

2) Hier wurden Produkt-Moment-Koeffizienten berechnet.

legungen müssen jedoch spekulativ bleiben, da die Autoren keine Angaben über die jeweils zugehörigen Varianzen der FAT-Rohwerte machen.

Für den Zusammenhang zwischen AN (ZD-Werte) und Intelligenz werden niedrige und in allen Fällen negative Koeffizienten gefunden (FAT/ZA: $r = -.1343$; FAT/ZS: $r = -.0391$; FAT/RW: $r = -.1401$). Zwar sind die Koeffizienten für FAT/ZA und FAT/RW signifikant von Null verschieden, doch ist dies angesichts des recht großen Stichprobenumfangs ($N = 185$) kaum verwunderlich. Auch erscheint es daher kaum vertretbar, die durchweg negativen Vorzeichen im Sinne einer komplementären Beziehung zwischen AN und Intelligenz zu deuten. Es kommt in bezug auf den verwendeten Intelligenztest hinzu, daß die Testleistung im FAT, einem eher auf Leistungsqualität ausgerichteten Verfahren, wohl weit weniger an motivationale Faktoren geknüpft ist, als dies bei einem auf Schnelligkeit angelegten Test zu erwarten wäre. Dieser Gesichtspunkt war bei der Auswahl des Tests bedacht worden, um nicht schon durch das verwendete Verfahren einen Zusammenhang zum AN nahezu legen (vgl. S. 55 f.).

Von den zur Beziehung zwischen AN, Intelligenz und Schulleistung vorliegenden Untersuchungen (siehe Punkt 2.4.2.) sind nur zwei (GOULD und KAPLAN 1940; HELM 1967) so angelegt, daß sie auch eine Aussage zum Zusammenhang zwischen AN und Intelligenz ermöglichen. - Bei den übrigen Untersuchungen wird die Variable Intelligenz zusammen mit der Schulleistung lediglich zur Vpn-Gruppierung herangezogen (SEARS 1940, 1941; SCHERER und SCHLIEP 1974) und in einem Fall (WALTER und MARZOLF 1951) gar nicht einbezogen. - GOULD und

KAPLAN (vgl. S. 23 ff.) geben für die Beziehung zwischen einem mittleren ZD-Wert (Median für 6 verschiedene AN-Aufgaben) und der Leistung in einem Intelligenztest ("Thorndike Test of Scholastic Ability for High School Graduates") einen Koeffizienten von $-.12^{1)}$ an. Allerdings bleibt zu bedenken, daß dieses Ergebnis sich auf die Daten einer studentischen Stichprobe bezieht; gleiches gilt für die Untersuchung von HELM. Nachdem HELM sich darauf beschränkt, die signifikanten Korrelationen anzugeben, und für die Beziehung zwischen Intelligenz (Verbal-IQ des HAWIE) und AN (ZD-Wert aus einer Zahlenadditionsaufgabe) kein Ergebnis mitgeteilt wird, ist daraus auf eine nur unbedeutende Korrelation zu schließen.

Die vor allem interessierende Beziehung ist die zwischen dem AN und der Schulleistung. Insgesamt gesehen haben sich für die Korrelationen zwischen den Kennwerten des ANs (ZD-Werte für die Verfahren ZA, ZS und RW) und der Schulleistung (verschiedene Notendurchschnitte) niedrige und zumeist (mit einer Ausnahme) negative Koeffizienten ergeben. (Die einzelnen Koeffizienten sind der Tabelle 17 zu entnehmen.) Zwar fallen einige der Koeffizienten (vorwiegend bei den Korrelationen der Notendurchschnitte mit den ZDn für ZA) signifikant aus, doch muß auch hier wiederum der bei der Signifikanzprüfung durchschlagende, große Stichprobenumfang ($N = 185$) berücksichtigt werden. Um - über eine statistische Signifikanz hinausgehend - auch von einer sogenannten praktischen Signifikanz sprechen zu können, müßten die Koeffizienten schon merklich höher ausfallen. Die in

1) Die Art der berechneten Korrelation wird nicht angegeben.

der vorliegenden Untersuchung gefundenen Interkorrelationen von Kennwerten des ANs und der Schulleistung sind daher nicht geeignet, eine (monotone) Beziehung zwischen diesen beiden Variablen annehmen zu lassen.

Es bleibt die Frage nach den Gründen für diese ausbleibende Beziehung. Als erstes wäre es theoretisch denkbar, daß das von uns erhobene AN keinen Indikator für motivationale Prozesse im Zusammenhang mit Leistung darstellt. Diese Möglichkeit kann aber wohl nach allem, was die im Verlauf der Versuche gemachten Beobachtungen ergeben haben, als sehr unwahrscheinlich ausgeschlossen werden (siehe unter Punkt 4.1.2. und 4.1.4.). Das AN (erfaßt über die ZD) zeigt eine relativ hohe intraindividuelle Konsistenz über die Zeit (siehe Punkt 4.3. und 4.3.1.) und auch über verschiedene Leistungen hinweg (siehe Punkt 4.2. und 4.2.1.).

Des weiteren wäre es möglich, daß das AN tatsächlich nicht mit der Schulleistung zusammenhängt. Dies kann heißen, daß das, was im AN erfaßt wird, sich nicht mit dem deckt, was etwa mit dem "Fleiß" des Schülers umschrieben wird. Eine nicht unwesentliche Rolle könnte auch die Dauer der zu erbringenden Leistung spielen. Während es im AN-Versuch um eine eher kurzfristige Leistung geht, bedarf es für die schulische Leistung eines "längeren Atems". Akzentuiert wird im einen Fall das Bei-der-Sache-Sein, im anderen Fall das Bei-der-Sache-Bleiben. (Aufschlußreiche Hinweise zu dieser Frage wären vielleicht dadurch zu gewinnen, daß man im AN-Versuch die Anzahl der Durchgänge nicht im vorhinein festlegt, sondern den Versuch so lange fortführt, bis die Vp ihn selbst beendet.) Als ein weiterer, auf die Leistung bezogener Gesichtspunkt ist folgendes

in Betracht zu ziehen: Das, was in der Schule als Leistung zählt, stellt sich dem Schüler immer nur indirekt, im Urteil des Lehrers dar. Dagegen ist im AN-Versuch die Rückmeldung über die erreichte Leistung viel direkter gegeben. Dies kann bedeuten, daß es der Vp in der AN-Versuchssituation eher möglich ist, ihre Leistungen und Leistungsvornahmen aufeinander zu beziehen. Schließlich könnte man fragen, in welchem Umfang die schulische Leistungssituation Spielraum für das AN des einzelnen bietet. Es wäre denkbar, daß dem Schüler in der Schule weniger Wahlmöglichkeiten für seine individuelle AN-Setzung offenstehen als im AN-Versuch, wobei dies eher für den unteren Leistungsbereich vermutet werden kann. Zwar sind auch in den AN-Versuchen Leistungsgrenzen eingeführt, doch ist (wie die Ergebnisse erkennen lassen) das Spektrum möglicher Leistungen, innerhalb dessen Ziele gesetzt werden können, wohl genügend weit bemessen, um Decken- bzw. Schlenneffekte zu vermeiden (vgl. dazu Punkt 4. auf S. 42, S. 49 f. sowie S. 105 f. und S. 120). Sollte diese Vermutung zutreffen, so würde sich die Frage stellen, inwieweit das Konzept des ANs für den Bereich der schulischen Leistung als angemessen gelten kann.

Ein weiterer Erklärungsansatz für die in der vorliegenden Untersuchung nicht gefundene Beziehung kann in der als Maß für das AN verwendeten Zieldiskrepanz gesucht werden. Es wäre möglich, daß die ZD unter bestimmten Bedingungen kein geeignetes Maß für das AN darstellt. Die angesichts der Heterogenität der untersuchten Leistungen relativ hohe intraindividuelle Konsistenz des ANs, genauer der ZD (vgl. Tabelle 13 und S. 109), spricht andererseits dafür, daß das ZD-Maß ge-

eignet ist, Wesentliches von dem der AN-Setzung zugrundeliegenden Motivationsgeschehen einzufangen. (Dies wäre ein Argument dafür, den gedanklichen Ansatz der ZD nicht aufzugeben.) Daß ein auf diese Weise erfaßter Motivationsanteil an einer Leistung jedoch in keiner Beziehung zur Schulleistung stehen soll - zumindest keine, die mit den hier eingesetzten statistischen Methoden aufzeigbar ist, läßt Zweifel an der Tauglichkeit des ZD-Maßes entstehen. Auf die Empfindlichkeit des ZD-Maßes gegenüber Ausreißerwerten, die bei insgesamt niedrigen ZDen relativ höher zu veranschlagen ist als bei hohen ZDen, wurde bereits hingewiesen (siehe S.119f.). Als ein gegenüber Extremwerten robusteres Maß würde sich statt der Mittelung z. B. der Median anbieten. Eine mögliche Weiterentwicklung der ZD kann darin gesehen werden, daß man sie in Relation zur erreichten Leistung betrachtet, und zwar ausgehend von folgender Überlegung: Eine hohe ZD bei einer Vp, die an sich schon hohe Leistungen erzielt, wird psychologisch etwas anderes bedeuten als eine gleich hohe ZD bei einer Vp mit einer niedrigeren Leistung; denn man kann wohl nicht annehmen, daß solche externen Bewertungssysteme der Leistung den Vpn verborgen blieben. Entsprechend kann man sicherlich auch davon ausgehen, daß ein individuelles Bezugssystem der Leistung vorhanden ist. Dies heißt, daß einzelne ZDen, die sich auf unterschiedlich hohe Leistungen innerhalb des individuellen Leistungsspektrums beziehen, auch unterschiedlich gewichtet werden müßten. Weitere Lösungen sind in Maßen zu suchen, die geeignet sind, den individuellen Verlauf der AN-Setzung zu beschreiben. (Z. B. wäre - in Abhängigkeit vom gewählten Wahrscheinlichkeitstheoretischen Modell - an eine Regressionsanalytische Betrachtung der ZD-Werte über die Versuchs-

durchgänge zu denken.) Weiterführend wäre sehr wahrscheinlich auch der Versuch, ein dem AN angemessenes wahrscheinlichkeitstheoretisches Modell - mit inhaltlich sinnvollen Hypothesen über die Abhängigkeit der Ereignisse (Worauf bezieht sich die Vp bei ihrer AN-Setzung?) - aufzustellen und in entsprechenden Versuchen zu prüfen. Zum Teil lassen sich derartige Hypothesen bereits aus dem vorliegenden Datenmaterial gewinnen, die zu verfolgen allerdings einer separaten Arbeit vorbehalten bleiben muß.

Die niedrigen Korrelationen zwischen AN und Schulleistung könnten ferner dadurch zustande gekommen sein, daß sich innerhalb der Gesamtstichprobe gegenläufige Zusammenhänge überlagern, die bei einer differentiellen Auswertung getrennt würden. Betrachtet man hierzu die - in binordinalen Kontingenztafeln - dargestellte Beziehung zwischen den ZD-Werten (getrennt für ZA, ZS und RW) und den Notendurchschnitten für das 1. Halbjahr (GN 1), so scheint diese Vermutung sich nicht zu bestätigen (siehe Tabelle 18 und dazu Abbildung 9 und 10). Auch ein (durchaus vorstellbarer) bitoner Zusammenhang - von der Art, daß bei guten Schülern vorwiegend niedrige, bei schlechten Schülern gehäuft hohe (positive und negative) ZDen auftreten - deutet sich nicht an (vgl. Tabelle 18). - Eine derartige Beziehung zeigte sich der Tendenz nach in dem von HELM (1967) mitgeteilten Ergebnis. HELM findet für die Korrelation zwischen |ZD|-Werten und Studienleistung (Notenmittel aller Hauptfächer) einen Koeffizienten von $r = -.37$ (vgl. S. 27 f.).

Zu dem gefundenen Ergebnis, daß die Schulleistung weit enger mit der Intelligenz als mit dem AN korreliert, ist noch folgendes anzumerken: Intelligenztests werden in der Regel auch an Schulnoten validiert, was für den in dieser Untersuchung verwendeten FAT 4-6 ebenfalls zutrifft. Damit ist bereits mit der Konstruktion des Intelligenztests, zumindest in gewissem Umfang, ein Zusammenhang mit der Schulleistung vorweggenommen.

Um festzustellen, inwieweit die gefundenen r -Korrelationen (Tabelle 17) dem vermittelnden Einfluß der jeweils dritten Variablen zuzuschreiben sind, wurden partielle r -Korrelationen bestimmt (siehe Tabelle 19). Die partiellen r -Koeffizienten liegen, wie zu erwarten, in allen Fällen (ausgenommen $r_{ZS/FAT.GN\ 1}$) leicht unterhalb der jeweils zugehörigen r -Koeffizienten (zum Vergleich siehe Tabelle 19).

Für den Zusammenhang der Schulleistung (GN 1) mit den beiden Variablen AN (ZD-Werte für ZA, ZS und RW) und Intelligenz (FAT-Rohwerte) haben sich multiple r -Koeffizienten zwischen .6306 und .6375 ergeben (Tabelle 20). Nimmt man die Korrelation zwischen Intelligenz und Schulleistung ($r_{FAT/GN\ 1} = .6296$) als Bezugspunkt, so fallen die multiplen Koeffizienten nur geringfügig höher aus, was auf Grund der jeweiligen Zweierbeziehungen zwischen den drei Variablen (siehe dazu Abbildung 11) bereits erwartet werden konnte.

In der mit Studenten durchgeführten Untersuchung von GOULD und KAPLAN (1940) ergeben sich für den Zusammenhang zwischen Gesamtdurchschnittsnoten und ZD-Werten (für 6 verschiedene AN-Aufgaben) Koeffizienten zwischen -.17 und .12 (siehe Tabelle 2, S. 25). Die Aussagefähigkeit dieser Ergebnisse bleibt allerdings, auf

Grund der bereits geschilderten Umstände (siehe S. 25 f.) unklar. Die weiteren Untersuchungen, die zur Beziehung zwischen AN, Intelligenz und Schulleistung dargestellt wurden (SEARS 1940, 1941; WALTER und MARZOLF 1951; SCHERER und SCHLIEP 1974), sind sämtlich so angelegt, daß Gruppen guter und schlechter Schüler im Hinblick auf das AN verglichen werden. SEARS findet bei guten Schülern ($N = 12$) im Durchschnitt niedrigere ZDen mit einer geringen Streuung, dagegen bei schlechten Schülern ($N = 12$) höhere ZDen mit einer höheren Streuung (vgl. Tabelle 1, S. 22). Die Gruppen waren nach der Intelligenzleistung parallelisiert. In die gleiche Richtung weisen die Ergebnisse von SCHERER und SCHLIEP. Bei Hauptschülern ($N = 22$) ergibt sich im Mittel eine kleinere Zieldifferenz ($|ZD|$ -Wert, relativiert auf die durchschnittliche Leistung) als bei Sonderschülern ($N = 11$) (siehe Tabelle 4, S. 30). Dieses Ergebnis weist darauf hin, daß das AN geeignet ist, zwischen guten und schlechten Schülern zu trennen. Allerdings muß bei diesem Ergebnis der sehr ungleiche Stichprobenumfang der verglichenen Gruppen berücksichtigt werden. Auch die nach der Intelligenz vorgenommene Parallelisierung beider Gruppen (siehe Tabelle 3, S. 29) birgt gewisse Schwierigkeiten, die die Interpretierbarkeit beeinträchtigen (vgl. S. 31). WALTER und MARZOLF (1951) stellen in ihrer Untersuchung fest, daß das AN (erfaßt über ZD-Werte) guter Schüler sich im Durchschnitt nicht bedeutsam von dem schlechter Schüler unterscheidet. Als Indikator der Schulleistung wurden bereits vorliegende Ergebnisse aus Schulleistungstests verwendet (siehe S. 26 f.). Die sich hier anschließende Frage, ob bzw. inwieweit ein Testdatum geeignet sein kann, die für den Schüler

erlebbare schulische Wirklichkeit abzubilden, wurde bereits diskutiert (siehe S. 27 und S. 59 f.).

Versucht man, zur Frage der Beziehung zwischen AN und Schulleistung ein abschließendes Résumé zu ziehen, so bleibt festzuhalten, daß die in dieser Untersuchung gewonnenen Ergebnisse keinen Hinweis auf einen solchen Zusammenhang ergeben haben. Von den an dieses Ergebnis angeknüpften Fragen läßt wohl vor allem die nach einem angemesseneren AN-Maß weiteren Aufschluß über die Frage der Beziehung von AN und Schulleistung erwarten.

5. Zusammenfassung

Die bisher geleistete Forschung über intellektuelle und motivationale Faktoren der Schulleistung wurde dargestellt. Gegenüber einer deutlichen Akzentuierung der intellektuellen Voraussetzungen ist die Frage nach den motivationalen Anteilen der Schulleistung (im engeren Sinn) erst sehr spät zum Gegenstand von Untersuchungen gemacht worden, was vor allem auf - nach wie vor nicht befriedigend gelöste - Schwierigkeiten der methodischen Umsetzung zurückgeführt werden kann.

In der vorliegenden Arbeit wurde versucht, über das Konzept des ANs einen möglichst phänomennahen Zugang zur Leistungsmotivation (verstanden im Sinne tatsächlich wirksamer Ansprüche an die eigene Leistung) zu finden.

Ein erstes Ziel dieser Untersuchung war es, die Beziehung zwischen den Variablen AN, Intelligenz und Schulleistung näher zu erkunden. Aus inhaltlichen und methodischen Gründen ergab sich daraus die vorab zu klärende Frage nach der intraindividuellen Konsistenz des ANs über verschiedene Leistungen (Generalität) und über die Zeit (Stabilität).

Die Daten zur Intelligenz (FAT 4-6), zum AN (drei verschiedene Verfahren: Zahlenaddition, Zahl-Symbol-Substitution, Ringewerfen) und zur Schulleistung (Noten im Zwischen- und im Jahreszeugnis) wurden an 185 Schülern des 5. Jahrgangs einer Gesamtschule erhoben. Zur Kontrolle der Stabilität des ANs fand nach einem Zeitintervall von 3 1/2 Monaten eine Versuchswiederholung der AN-Verfahren statt.

Das Material aus den AN-Versuchen gibt deutliche Hinweise, daß die Erfassung des ANs unter den hier gegebenen Versuchsbedingungen in geeigneter Weise möglich ist. Die Gesamtmittelwerte der in den einzelnen Verfahren jeweils vorgenommenen und erreichten Leistungen zeigen sowohl für die erste wie für die wiederholte Versuchsdurchführung an, daß es in recht guter Annäherung gelungen ist, einen individuell angemessenen mittleren Schwierigkeitsbereich einzuführen. Hierin liegt eine wichtige Voraussetzung für das Zustandekommen echter AN-Bildungen und für eine hohe Zuverlässigkeit.

Die zur Frage der Generalität des ANs ermittelten Interkorrelationen der ZD-Werte aus den verschiedenen AN-Verfahren weisen in allen Fällen auf einen mittleren Grad des Zusammenhangs hin ($ZA/ZS: r_s = .6925$; $ZA/RW: r_s = .5876$; $ZS/RW: r_s = .5889$). Berücksichtigt man die Heterogenität der miteinander verglichenen Leistungen - motorische Fertigkeiten, kognitive Fähigkeiten -, so kann das Maß an intraindividueller Konsistenz als relativ hoch angesehen werden. Die Verfahren ZA und ZS korrelieren untereinander etwas höher als jedes der beiden Verfahren mit RW. Mögliche Gründe für diese Abweichung (z. B. Art der abverlangten Leistung, Leistungsverlauf, Anschaulichkeit der Leistungssituation) werden diskutiert.

Die - ein relativ langes Zeitintervall umfassende - Stabilitätskontrolle des (über ZD-Werte erfaßten) ANs erbringt klare Hinweise auf ein relativ hohes Maß an intraindividueller Konsistenz. Man darf annehmen, daß der gewählte Zeitabstand genügend groß war, um einerseits Erinnerungseinflüsse weitgehend ausschließen zu können und andererseits etwaige Merkmalsschwankungen

zur Wirkung gelangen zu lassen. Die Stabilität fällt - wie zu erwarten - für das auf allen drei AN-Verfahren basierende Gesamtmaß ($\overline{ZD}_{rel.}$) am höchsten aus ($r_s = .7780$). Ein Koeffizient dieser Höhe würde bereits den Ansprüchen genügen, die derzeit an die Retest reliabilität standardisierter Tests gestellt werden.

Über die Beziehung von AN, Intelligenz und Schulleistung geben die zwischen den Kennwerten dieser drei Variablen ermittelten Korrelationen Aufschluß. Die für den Zusammenhang zwischen Intelligenz (FAT-Rohwerte) und Schulleistung (verschiedene Notendurchschnitte) gefundenen Korrelationen (r -Koeffizienten) bewegen sich zwischen .4492 und .6503. Die Koeffizienten entsprechen in ihrer Größenordnung den in der Literatur berichteten Ergebnissen. Die Korrelationen zwischen AN (ZD-Werte für die drei AN-Verfahren) und Intelligenz lassen auf einen nur unbedeutenden Zusammenhang zwischen beiden Variablen schließen. Die niedrigen und durchweg negativen Koeffizienten fallen zwar für FAT/ZA und FAT/RW signifikant aus, doch ist dies angesichts des recht großen Stichprobenumfangs ($N = 185$) kaum verwunderlich. Die für die (vor allem interessierende) Beziehung von AN (ZD-Werte für ZA, ZS und RW) und Schulleistung (verschiedene Notendurchschnitte) erhaltenen Korrelationen (r -Koeffizienten) geben keinen Hinweis auf einen nennenswerten Zusammenhang zwischen diesen beiden Variablen. Es werden niedrige und zumeist (mit einer Ausnahme) negative Koeffizienten gefunden. Zwar werden einige der Koeffizienten signifikant, doch muß auch hier der bei der Signifikanzprüfung durchschlagende, große Stichprobenumfang berücksichtigt werden. Eine differentielle Betrachtung läßt ebenfalls keine Beziehung erkennen. Auch ein bitoner Zusammenhang zeichnet sich nicht ab.

An dieses Ergebnis einer ausbleibenden Beziehung werden Fragen angeknüpft, die einer weiteren Klärung bedürfen. Von den hier aufgeworfenen Fragen läßt wohl vor allem die nach einem angemesseneren Maß für das AN weiteren Aufschluß über die Beziehung zwischen AN und Schulleistung erwarten. Es werden Möglichkeiten vorgeschlagen, die sich auf Weiterentwicklungen des ZD-Maßes und auf neue Ansätze zum AN-Maß beziehen. Hier müssen sich weitere Untersuchungen anschließen.

LITERATUR

- ARNOLD, W.: Der Pauli-Test. München 1970 (erschienen als 4., verbesserte u. erweiterte Aufl. von R. PAULI: Der Arbeitsversuch als charakterologisches Prüfverfahren. Leipzig 1943.).
- BELSER, H.; ANGER, H.; BARGMANN, R.; RAATZ, U.: Frankfurter Analogietest für 4. bis 6. Klassen. FAT 4-6. Reihe Deutsche Schultests. Hrsg. v. Dt. Inst. f. Intern. Päd. Forsch. Weinheim 1973.
- BELSER, H.; RAATZ, U.: Frankfurter Analogietest. FAT 4-6. Reihe Deutsche Schultests. Hrsg. v. Dt. Inst. f. Intern. Päd. Forsch. Beiheft mit Anleitung und Normmentabellen. Bearbeitung und Normierung 1972. Weinheim 1973.
- BONDY, C. (Hrsg.): Hamburg-Wechsler-Intelligenztest für Erwachsene. 3. unveränd. Aufl. Bern 1964.
- BRUCKS, Christel-A.: Eine Untersuchung zur Stabilität des Anspruchsniveaus. In: Untersuchungen zum Forschungsschwerpunkt "Hauptschulprobleme". Fachbereich Erziehungswissenschaften der Universität Bayreuth. 1977, 29-48.
- DÜHRSEN, Annemarie: Psychogene Erkrankungen bei Kindern und Jugendlichen. Göttingen 1954.
- DÜKER, H.: Untersuchungen über die Ausbildung des Wollens. Bern, Stuttgart, Wien 1975.
- FESTINGER, L.: Wish, expectation, and group standards as factors influencing level of aspiration. J. abnorm. soc. Psychol., 1942, 37, 184-200.
- FIPPINGER, F.: Intelligenz und Schulleistung. Eine experimentelle Untersuchung bei 9- bis 10jährigen Schülern. München 1966.
- FRANK, J. D.: Individual differences in certain aspects of the level of aspiration. Amer. J. Psychol., 1935, 47, 119-128.
- GARDNER, J. W.: Level of aspiration in response to a prearranged sequence of scores. J. exp. Psychol., 1939, 25, 601-621.

- GOULD, Rosalind: Factors underlying expressed "level of aspiration". J. Psychol., 1938, 6, 265-279.
- GOULD, Rosalind: An experimental analysis of "level of aspiration". Genetic Psychology Monographs, 1939, 21, 3-115.
- GOULD, Rosalind; KAPLAN, N.: The relationship of "level of aspiration" to academic and personality factors. J. soc. Psychol., 1940, 11, 31-40.
- HEATHERS, Louise B.: Factors producing generality in the level of aspiration. J. exp. Psychol., 1942, 30, 392-406.
- HECKHAUSEN, H.: Motivationsanalyse der Anspruchsniveausetzung. Psychol. Forsch., 1955, 25, 118-154.
- HECKHAUSEN, H.: Hoffnung und Furcht in der Leistungsmotivation. Meisenheim 1963. (a)
- HECKHAUSEN, H.: Eine Rahmentheorie der Motivation in zehn Thesen. (Tübinger Symposion über Motivation). Z. exp. angew. Psychol., 1963, 10, 604-626. (b)
- HECKHAUSEN, H.: Leistungsmotivation. In: H. THOMAE (Hrsg.): Handbuch der Psychologie. Bd. 2. Göttingen 1965, 602-702.
- HECKHAUSEN, H.: The anatomy of achievement motivation. New York 1967.
- HECKHAUSEN, H.: Motivation in der Leistungsgesellschaft. Die Deutsche Schule, 1968, 10, 637-648.
- HECKHAUSEN, H.: Motivation: Kognitionspsychologische Aufspaltung eines summarischen Konstrukts. Psychol. Rundschau, 1977, 28, 175-189.
- HELM, J.: Studien zur Korrelationsanalyse von Variablen der Leistungsmotivation. Z. exp. angew. Psychol., 1967, 14, 600-615.
- HÖRMANN, H.: Theoretische Grundlagen der projektiven Tests. In: R. HEISS (Hrsg.): Handbuch der Psychologie. Bd. 6. Göttingen 1964, 71-144.
- HOPPE, F.: Erfolg und Mißerfolg. Psychol. Forsch., 1930, 14, 1-62.

- INGENKAMP, K.: Pädagogische Diagnostik. Forschungsbericht über Schülerbeurteilung in Europa. Weinheim, Basel 1975.
- INGENKAMP, K. (Hrsg.): Die Fragwürdigkeit der Zensurengebung. 6. überarbeitete u. erweiterte Aufl. Weinheim 1976.
- JÄGER, A. O.: Einige emotionale, conative und zeitliche Bedingungen des Erinnerns. Diss. Göttingen 1958.
- JUCKNAT, Margarete: Leistung, Anspruchsniveau und Selbstbewußtsein. Psychol. Forsch., 1937, 22, 89-179.
- KEMMLER, Lilly; LANGHEINRICH, D.: Eine deutsche Fassung der "Primary Mental Abilities". Ber. 25. Kongreß d. Dt. Ges. f. Psychol. in Münster 1966. Göttingen 1967, 669-678.
- KENDALL, M. G.: Rank correlation methods. 4th edit. London 1970.
- KLINGER, E.: Fantasy need achievement as a motivational construct. Psychol. Bull., 1966, 66, 291-308.
- KORNADT, H.-J.: Thematische Apperzeptionsverfahren. In: R. HEISS (Hrsg.): Handbuch der Psychologie. Bd. 6. Göttingen 1964, 635-684.
- KOWALEWSKY, W.: Die Wirkungen verschiedener Erfolgs- und Mißerfolgssituationen auf das Anspruchsniveau. Diss. Göttingen 1955.
- KRÜGER, H.-P.: Erlebte Interaktionshäufigkeiten als soziometrische Maße. Diss. Erlangen 1973.
- KRÜGER, H.-P.: Soziometrie in der Schule. Weinheim, Basel 1976.
- KRUMBOLTZ, J. D.: Measuring achievement motivation - a review. J. Counseling Psychol., 1957, 4, 191-198.
- LANGFELDT, H. P.; FINGERHUT, W.: Empirische Ansätze zur Aufklärung des Konstruktes "Schulleistung". In K. HELLER (Hrsg.): Leistungsbeurteilung in der Schule. Heidelberg 1974, 38-45.
- LAVIN, D. E.: The prediction of academic performance. New York 1965.

- LEWIN, K. (Hrsg.): Untersuchungen zur Handlungs- und Affektpsychologie. Psychol. Forsch., 1926-1934.
- LEWIN, K.; DEMBO, Tamara; FESTINGER, L.; SEARS, Pauline S.: Level of aspiration. In: J. McV. HUNT (Hrsg.): Personality and the behavior disorders. Bd. 1. New York 1944, 333-378.
- LIENERT, G. A.: Testaufbau und Testanalyse. 2. durchges. u. verb. Aufl. Weinheim, Berlin 1967.
- LIENERT, G. A.: Verteilungsfreie Methoden in der Biostatistik. 2. völlig neu bearb. Aufl. Meisenheim 1973.
- LÖSCHENKOHL, E.: Gibt es einen allgemein faßbaren Zusammenhang zwischen Schulleistung und Intelligenz? Psychol. in Erzieh. u. Unterricht, 1973, 20, 145-155.
- McCLELLAND, D. C.: Methods of measuring human motivation. In: J. W. ATKINSON (Hrsg.): Motives in fantasy, action, and society. Princeton, N. J. 1958, 7-42.
- McCLELLAND, D. C.: Assessing human motivation. New York 1971.
- McCLELLAND, D. C.; ATKINSON, J. W.; CLARK, R. A.; LOWELL, E. L.: The achievement motive. New York 1953.
- MEIS, R.: Diagnostischer Rechtschreibtest für 4. und 5. Klassen. DRT 4-5. Reihe Deutsche Schultests. Hrsg. v. K. INGENKAMP. Weinheim, Berlin 1970.
- MEYER, W.-U.; HECKHAUSEN, H.; KEMMLER, Lilly: Validierungskorrelate der inhaltsanalytisch erfaßten Leistungsmotivation guter und schwacher Schüler. Psychol. Forsch., 1965, 28, 301-328.
- MÜLLER, R.: Zur "neurotischen Intelligenzhemmung" bei Schulkindern. Praxis d. Kinderpsychol. u. Kinderpsychiat., 1955, 4, 297-299.
- MURRAY, H. A.: Explorations in personality. New York 1938.
- MURRAY, H. A.: Thematic apperception test. Cambridge 1942.
- PIECHOCKI, H.: Lernhemmungen bei neurotischer Fehlentwicklung. Praxis d. Kinderpsychol. u. Kinderpsychiat., 1952, 1, 11-17.

- PRESTON, M. G.; BAYTON, J. A.: Correlations between levels of aspiration. J. Psychol, 1942, 13, 369-373.
- ROTTER, J. B.: Level of aspiration as a method of studying personality. II. Development and evaluation of a controlled method. J. exp. Psychol., 1942, 31, 410-422.
- SACHS, L.: Angewandte Statistik. Berlin, Heidelberg, New York 1974. (Zugleich 4. Neubearb. u. erw. Aufl. der "Statistischen Auswertungsmethoden" mit neuer Bibliographie).
- SCHERER, J.; SCHLIEP, Monika: Persönlichkeitsmerkmale und Leistungsverhalten bei gleichintelligenten Haupt- und Sonderschülern des 5. Schuljahres. Psychol. in Erzieh. u. Unterricht, 1974, 21, 81-90.
- SCHMALOHR, E.: Gruppennormen des "HAWIK" im Hilfsschul-Überweisungsverfahren. Zeitschr. f. Heilpäd., 1962, 13, 165-176.
- SCHMALT, H.-D.: Methoden der Leistungsmotivmessung. In: H.-D. SCHMALT und W.-U. MEYER (Hrsg.): Leistungsmotivation und Verhalten. Stuttgart 1976, 165-191. (a)
- SCHMALT, H.-D.: Die Messung des Leistungsmotivs. Göttingen, Toronto, Zürich 1976. (b)
- SEARS, Pauline S.: Levels of aspiration in academically successful and unsuccessful children. J. abnorm. soc. Psychol., 1940, 35, 498-536.
- SEARS, Pauline S.: Level of aspiration in relation to some variables of personality: clinical studies. J. soc. Psychol., 1941, 14, 311-336.
- STAABS, Gerdhild v.: Schulschwierigkeiten in tiefenpsychologischer Sicht. Praxis d. Kinderpsychol. u. Kinderpsychiat., 1955, 4, 40-45.
- STERN, E.: Deblilität, Demenz, Pseudodeblilität, Intelligenzhemmung und Schulleistung. Schule u. Psychol. 1956, 3, 364-379.
- TRAXEL, W.: Empirische Untersuchungen zur Einteilung der Gefühlsqualitäten. Ber. 22. Kongreß d. Dt. Ges. f. Psychol. in Heidelberg 1959. Göttingen 1960, 290-294.

TRAXEL, W.: Über Dimensionen und Dynamik der Motivierung.
Z. exp. angew. Psychol., 1961, 8, 418-428.

TRAXEL, W.: Grundzüge eines Systems der Motivierungen.
Arch. ges. Psychol., 1962, 114, 143-172.

TRAXEL, W.: Ein Erkundungsexperiment zum System der Motivierungen. Psychol. Beiträge, 1966, 9, 309-322.

WALTER, L. M.; MARZOLF, S. S.: The relation of sex, age and school achievement to levels of aspiration.
J. educ. Psychol., 1951, 42, 285-292.

WASNA, Maria: Motivation, Intelligenz und Lernerfolg.
2. durchges. Aufl. München 1976.

ZULLIGER, H.: Untersuchung eines Falles von Pseudo-
Debilität. Psyche, 1957, 11, 388-400.

A N H A N G

	Seite
Versuchsmaterial für die AN-Verfahren	
Zahlenaddition	1 - 3
Zahl-Symbol-Substitution	4 - 16
Ringewerfen	17 - 18
(Registrierbogen mit den Daten einer Vp)	
Tabelle 1	19

Name: _____

Vorname: _____

Alter: _____

Klasse: _____

9	5	6	7	3	7	5	7	2	9	4
7	7	5	6	2	8	5	6	1	6	3
6	6	8	8	6	6	3	3	6	7	6
5	8	7	5	4	4	7	5	4	3	2
3	3	3	3	8	5	6	4	3	8	7
7	1	1	8	7	8	2	2	2	5	6
6	6	5	6	2	9	1	9	9	9	9
3	5	3	2	6	5	9	8	7	2	6
2	3	9	9	5	7	2	5	8	1	8
1	9	6	4	9	4	8	2	4	5	7
6	7	7	7	7	3	4	7	7	3	2
9	8	8	5	8	7	7	6	5	7	6
8	4	4	3	5	5	6	3	9	6	4
4	3	2	5	6	9	1	2	3	4	7
5	6	8	3	3	4	4	5	7	1	5
2	5	7	8	2	7	3	4	6	3	8
3	7	4	7	5	6	7	8	5	2	3
1	2	6	3	8	2	6	7	3	7	7
7	1	5	5	7	8	9	6	8	6	6
8	8	9	4	6	7	4	2	6	4	5
6	4	7	9	3	2	2	9	3	8	9
9	7	6	5	1	6	6	7	2	6	4
5	9	4	8	5	4	5	2	5	2	2
3	8	6	1	4	7	1	1	4	6	5
4	3	5	7	8	8	2	7	1	5	7
8	2	2	5	7	9	7	5	9	1	8
7	7	3	6	5	3	5	6	8	4	5
4	5	1	3	6	9	6	3	4	9	3
5	6	5	6	4	4	3	2	7	8	2
8	4	6	2	5	8	2	9	8	3	9



Mit dem, was ich gerade geschafft habe, bin ich:



= sehr zufrieden



= zufrieden



= unzufrieden



= sehr unzufrieden



= weiß nicht

1	=
---	---

2	×
---	---

3	>
---	---

4	○
---	---

5	□
---	---

6	└
---	---

7	□
---	---

8	—
---	---

3	1	7	4	5	2	7	1	6	8	2	5	8	3	6	2	1	7	5	6	4	8	3	4	



1	=
---	---

2	×
---	---

3	<
---	---

4	○
---	---

5	□
---	---

6	└
---	---

7	□
---	---

8	-
---	---

--

5	2	7	6	1	4	8	3	1	5	8	2	1	4	7	3	4	6	5	3	8	6	2	7

--



1	=
---	---

2	×
---	---

3	>
---	---

4	○
---	---

5	□
---	---

6	└
---	---

7	□
---	---

8	-
---	---

--

1	7	3	2	8	5	3	6	4	7	8	5	1	4	7	8	6	2	5	1	3	2	6	4	

--





8	—
---	---

7	□
---	---

6	└
---	---

5	┐
---	---

4	○
---	---

3	<
---	---

2	×
---	---

1	=
---	---

7	3	8	5	1	6	2	4	7	5	1	2	5	8	2	1	4	3	6	4	7	3	8	6	

--

1	=
---	---

2	×
---	---

3	>
---	---

4	○
---	---

5	□
---	---

6	└
---	---

7	□
---	---

8	-
---	---

--

4	7	2	5	8	7	1	4	3	6	2	1	6	3	8	5	1	3	2	8	5	7	4	6

--



1	=
---	---

2	×
---	---

3	>
---	---

4	○
---	---

5	□
---	---

6	└
---	---

7	□
---	---

8	—
---	---

--

3	2	6	1	8	4	7	2	6	5	1	3	5	8	7	2	4	3	8	5	1	6	4	7	

--



1	=
---	---

2	×
---	---

3	>
---	---

4	○
---	---

5	□
---	---

6	└
---	---

7	□
---	---

8	-
---	---

--

8	5	2	4	1	3	7	5	8	6	3	7	5	4	1	6	2	7	8	3	2	1	4	6	

--



1	=
---	---

2	×
---	---

3	>
---	---

4	○
---	---

5	□
---	---

6	└
---	---

7	⊂
---	---

8	-
---	---

--

2	1	3	7	5	4	8	6	2	7	3	1	4	8	6	3	5	8	4	1	2	6	7	5

--



1	=
---	---

2	×
---	---

3	∧
---	---

4	○
---	---

5	□
---	---

6	└
---	---

7	□
---	---

8	—
---	---

6	7	3	8	5	1	2	4	1	3	7	2	8	5	1	6	8	4	2	7	4	5	6	3



1	=
---	---

2	×
---	---

3	∧
---	---

4	○
---	---

5	□
---	---

6	└
---	---

7	□
---	---

8	-
---	---

7	6	3	5	1	2	7	3	8	4	1	5	4	8	6	7	4	2	8	3	6	1	2	5



Mit dem, was ich gerade geschafft habe, bin ich:

 = sehr zufrieden

 = zufrieden

 = unzufrieden

 = sehr unzufrieden

 = weiß nicht

Name: K.Vorname: KweAlter: 11Klasse: 5.6Datum: 12.7.1976Beginn: 9.23Ende: 9.39Vp-Stab:
1,00
1,05

Durchgänge	AN	LN	++/+/--/0	
0	-	6		
1	6	11	++	zu AN=6: "überschätzen soll man doch nicht!"
2	8	10	+	
3	9	12	++	
4	10	12	++	
5	12	10	-	
6	10	13	++	
7	11	12	++	
8	12	16	++	
9	12	10	-	
10	10	11	+	
11	11	13	+	

Bemerkungen:

Mit dem, was ich gerade geschafft habe, bin ich:

sehr zufrieden

zufrieden

unzufrieden

sehr unzufrieden

weiß nicht

Tabelle 1: Mittelwerte und Standardabweichungen der AN- und LN-Werte pro Durchgang über alle Vpn (N = 185), getrennt für die Verfahren ZA, ZS und RW.

AN-Verfahren	Durchgang	AN		LN	
		$\bar{X}$	s	$\bar{X}$	s
ZA	1	10,497	4,437	11,465	4,117
	2	11,449	3,938	11,930	3,962
	3	11,822	3,921	12,805	3,990
	4	12,270	4,003	12,362	4,384
	5	12,335	4,080	13,757	4,416
	6	12,595	4,399	12,714	4,418
	7	12,714	4,358	16,135	4,068
	8	13,551	4,644	15,924	4,528
	9	14,254	4,744	15,016	4,653
	10	13,524	4,863	16,292	5,099
	11	14,632	5,294	15,151	4,755
ZS	1	8,232	3,201	9,005	2,744
	2	8,822	2,931	10,681	2,886
	3	9,789	3,251	11,032	2,690
	4	10,211	3,015	11,697	2,833
	5	10,697	3,168	12,978	2,965
	6	11,503	3,281	12,600	2,855
	7	11,632	3,276	12,989	3,027
	8	11,849	3,421	13,578	3,201
	9	12,168	3,514	14,260	3,523
	10	12,541	3,714	14,270	3,212
	11	13,081	3,892	14,622	3,805
RW	1	7,081	2,670	7,076	2,683
	2	7,497	2,588	7,465	2,590
	3	7,919	2,716	7,924	2,822
	4	8,222	2,748	7,930	2,910
	5	8,308	2,777	7,838	2,643
	6	8,151	2,666	8,141	2,969
	7	8,368	2,698	8,173	2,810
	8	8,384	2,811	7,768	2,669
	9	8,049	2,677	8,265	2,633
	10	8,519	2,749	7,989	2,750
	11	8,389	2,947	8,157	2,659

Lebenslauf

Am 21. Mai 1946 wurde ich in Alfeld (Leine) geboren.

Von 1953 bis 1957 besuchte ich die Volksschule, dann das Gymnasium in Alfeld, wo ich 1966 das Abitur ablegte.

Zum Wintersemester 1966/67 schrieb ich mich an der Universität Kiel für das Fach Psychologie ein und absolvierte dort 1969 die Diplom-Vorprüfung für Psychologen. Im Wintersemester 1969/70 studierte ich an der Universität Wien. Zum Sommersemester 1970 kehrte ich an die Universität Kiel zurück, wo ich 1971 die Diplom-Hauptprüfung für Psychologen ablegte.

Von September 1971 bis Dezember 1972 war ich in Kiel als wissenschaftliche Angestellte beim schleswig-holsteinischen Kultusministerium tätig.

Seit dem 1. 7. 1973 bin ich mit der Verwaltung der Dienstgeschäfte eines wissenschaftlichen Assistenten beauftragt, zunächst an der Forschungsstelle für Hauptschulprobleme an der Zweiten Erziehungswissenschaftlichen Fakultät der Universität Erlangen-Nürnberg in Bayreuth, ab 1976 beim Lehrstuhl für Psychologie der Universität Bayreuth.

Meine akademischen Lehrer waren die Herren Professoren: K.-O. Apel, W. Bröcker, D. Dörner, G. Fischer, G. Guttmann, W. Hallermann, H. W. Jürgens, M. Monjé, J. Schaeuble, H. Schmitz, G.-E. Störring, H. Strotzka, W. Sünel, W. Toman, W. Traxel, D. Trincker, H. Völkel, C. Vogel, H. Wegener, H. Werbik, H. H. Wieck.

**Dissertationsdruck
GmbH & Co. KG**

M. SCHADEL

86 Bamberg · Hohe-Kreuz-Str.9b · Tel. 0951/5 44 14